BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

NCBI *results of* **BLAST**

BLASTP 2.2.7 [Jan-02-2004]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

RID: 1073935785-25692-185772007425.BLASTQ4

Query=
(398 letters)

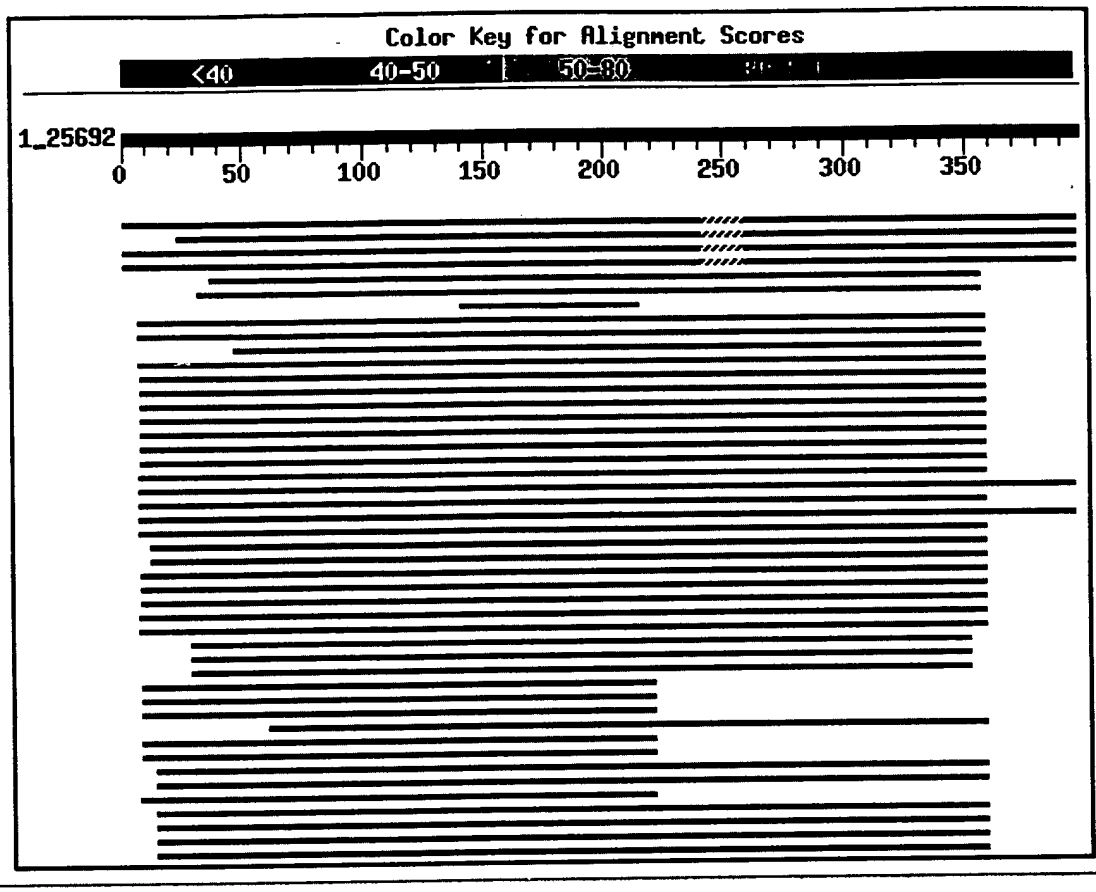
Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF
1,591,231 sequences; 521,258,069 total letters

If you have any problems or questions with the results of this search please refer to the **BLAST FAQs**

Taxonomy reports

Distribution of 50 Blast Hits on the Query Sequence

Mouse-over to show defline and scores. Click to show alignments



Sequences producing significant alignments:

	Score (bits)	E Value	
gi|16876435|ref|NP_473362.1| G protein-coupled receptor 101...	495	e-139	L
gi|21929156|dbj|BAC06152.1| seven transmembrane helix recep...	447	e-124	
gi|38086312|ref|XP_141764.3| similar to G protein-coupled r...	404	e-111	L
gi|34881783|ref|XP_229186.2| similar to G protein-coupled r...	404	e-111	L
gi|24476016|ref|NP_722561.1| G protein-coupled receptor 161...	137	3e-31	L
gi|6677701|ref|NP_031395.1| G protein-coupled receptor 161;...	134	4e-30	L
gi|29611554|gb|AAO85088.1| G protein-coupled receptor GPR10...	129	9e-29	
gi|8843925|gb|AAF80168.1| alpha 1a-adrenoceptor isoform 2 [...	124	3e-27	
gi|3023219|sp|O02824|A1AA_RABIT Alpha-1A adrenergic recepto...	124	3e-27	
gi|26788019|emb|CAC94897.2| SI:bZ20I5.4 (novel protein simi...	124	3e-27	
gi|8843927|gb|AAF80169.1| alpha 1a-adrenoceptor isoform 3 [...	124	4e-27	
gi|15004694|gb|AAK77197.1| adrenergic receptor alpha-1a [Ho...	122	9e-27	L
gi|15451759|ref|NP_150646.1| alpha-1A-adrenergic receptor i...	122	2e-26	L
gi|15451761|ref|NP_150647.1| alpha-1A-adrenergic receptor i...	121	2e-26	L
gi|4501961|ref|NP_000671.1| alpha-1A-adrenergic receptor is...	120	4e-26	L
gi|666893|gb|AAB59486.1| alpha-1C-adrenergic receptor	120	4e-26	L
gi|15451757|ref|NP_150645.1| alpha-1A-adrenergic receptor i...	120	5e-26	L
gi|1168246|sp|P35348|A1AA_HUMAN Alpha-1A adrenergic recepto...	120	5e-26	L
gi|1168247|sp|P43140|A1AA_RAT Alpha-1A adrenergic receptor ...	120	5e-26	L
gi|31542114|ref|NP_038489.2| adrenergic receptor, alpha 1a;...	119	7e-26	L

gi|20141255|sp|P97718|A1AA_MOUSE Alpha-1A adrenergic recept...	119	1e-25	L
gi|26351717|dbj|BAC39495.1| unnamed protein product [Mus mu...	119	1e-25	L
gi|2494939|sp|Q91175|A1AA_ORYLA Alpha-1A adrenergic recepto...	118	2e-25	
gi|7441613|pir||S71323 alpha-1A adrenergic receptor - Japan...	118	2e-25	
gi|8392870|ref|NP_058887.1| adrenergic receptor, alpha 1a; ...	118	2e-25	L
gi|409029|gb|AAA93114.1| alpha1C adrenergic receptor	117	3e-25	L
gi|547222|gb|AAB31165.1| alpha adrenergic receptor subtype ...	117	4e-25	
gi|14916519|sp|Q9WU25|A1AA_CAVPO Alpha-1A adrenergic recept...	117	4e-25	
gi|27806213|ref|NP_776923.1| adrenergic, alpha 1A, receptor...	116	8e-25	L
gi|6563386|emb|CAB62570.1| alpha-1A adrenergic receptor [Su...	115	1e-24	
gi|4501957|ref|NP_000669.1| alpha-1D-adrenergic receptor; a...	114	4e-24	L
gi|7690135|gb|AAB31163.2| alpha adrenergic receptor subtype...	113	5e-24	L
gi|86790|pir||JH0447 alpha-1A-adrenergic receptor - human >...	110	3e-23	L
gi|40362763|gb|AAR84650.1| alpha 1A adrenoceptor isoform 6 ...	110	5e-23	
gi|40362753|gb|AAR84645.1| alpha 1A adrenoceptor isoform 2c...	108	1e-22	L
gi|639573|gb|AAB30835.1| alpha 1c-adrenoceptor, alpha 1c-AR...	108	2e-22	
gi|40362757|gb|AAR84647.1| alpha 1A adrenoceptor isoform 3c...	108	2e-22	L
gi|40362751|gb|AAR84644.1| alpha 1A adrenoceptor isoform 2b...	108	2e-22	L
gi|40362759|gb|AAR84648.1| alpha 1A adrenoceptor isoform 5a...	107	3e-22	
gi|478273|pir||JC1525 alpha-1B-adrenergic receptor - rat >g...	107	5e-22	L
gi|8392867|ref|NP_058687.1| adrenergic receptor, alpha 1b; ...	106	8e-22	L
gi|5902693|sp|O77621|A1AA_CANFA Alpha-1A adrenergic recepto...	105	1e-21	
gi|1168245|sp|P35368|A1AB_HUMAN Alpha-1B adrenergic recepto...	105	2e-21	L
gi|112871|sp|P18841|A1AB_MESAU Alpha-1B adrenergic receptor...	105	2e-21	
gi|4501959|ref|NP_000670.1| alpha-1B-adrenergic receptor; a...	105	2e-21	L
gi|547221|gb|AAB31164.1| alpha adrenergic receptor subtype ...	105	2e-21	

Alignments

>gi|16876435|ref|NP_473362.1| L G protein-coupled receptor 101 [Homo sapiens]
gi|16566341|gb|AAL26486.1| L G protein-coupled receptor [Homo sapiens]
gi|20152240|dbj|BAB89301.1| putative G-protein coupled receptor [Homo sapiens]
Length = 508

Score = 495 bits (1274), Expect = e-139
Identities = 242/242 (100%), Positives = 242/242 (100%)

```
Query: 1   MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKP 60
            MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKP
Sbjct: 1   MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKP 60

Query: 61  QLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVN 120
            QLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVN
Sbjct: 61  QLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVN 120

Query: 121 TIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNA 180
            TIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNA
Sbjct: 121 TIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNA 180

Query: 181 LCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKD 240
            LCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKD
Sbjct: 181 LCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKD 240
```

```
Query: 241 CV 242
           CV
Sbjct: 241 CV 242
```

```
 Score =  231 bits (589), Expect = 2e-59
 Identities = 109/139 (78%), Positives = 110/139 (79%)

Query: 260 MNIPEXXXXXXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET 319
           +NIPE                CYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET
Sbjct: 370 VNIPESLPPSRRNSNSNPPLPRCYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET 429

Query: 320 QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDMLXXXXXXXXXXXXDSHPDLPG 379
           QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDML            DSHPDLPG
Sbjct: 430 QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDMLKKFFCKEKPPKEDSHPDLPG 489

Query: 380 TEGGTEGKIVPSYDSATFP 398
           TEGGTEGKIVPSYDSATFP
Sbjct: 490 TEGGTEGKIVPSYDSATFP 508
```

>gi|21929156|dbj|BAC06152.1| seven transmembrane helix receptor [Homo sapiens]

```
          Length = 485

 Score =  447 bits (1150), Expect = e-124
 Identities = 219/219 (100%), Positives = 219/219 (100%)

Query: 24  MPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISL 83
           MPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISL
Sbjct: 1   MPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISL 60

Query: 84  VAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVSVDRYLSIIHPLSYPSKM 143
           VAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVSVDRYLSIIHPLSYPSKM
Sbjct: 61  VAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVSVDRYLSIIHPLSYPSKM 120

Query: 144 TQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIWGASPSYTILSVVSFIVI 203
           TQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIWGASPSYTILSVVSFIVI
Sbjct: 121 TQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIWGASPSYTILSVVSFIVI 180

Query: 204 PLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKDCV 242
           PLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKDCV
Sbjct: 181 PLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKDCV 219
```

```
 Score =  231 bits (590), Expect = 1e-59
 Identities = 109/139 (78%), Positives = 110/139 (79%)

Query: 260 MNIPEXXXXXXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET 319
           +NIPE                CYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET
Sbjct: 347 VNIPESLPPSRRNSNSNPPLPRCYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET 406

Query: 320 QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDMLXXXXXXXXXXXXDSHPDLPG 379
           QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDML            DSHPDLPG
Sbjct: 407 QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDMLKKFFCKEKPPKEDSHPDLPG 466

Query: 380 TEGGTEGKIVPSYDSATFP 398
```

```
          TEGGTEGKIVPSYDSATFP
Sbjct: 467 TEGGTEGKIVPSYDSATFP 485
```

>gi|38086312|ref|XP_141764.3| similar to G protein-coupled receptor 101 [Mus mus
Length = 615

```
 Score =  404 bits (1037), Expect = e-111
 Identities = 189/242 (78%), Positives = 212/242 (87%)

Query: 1   MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKP 60
           M  +CTNST+E+N S  C+PLSKMPIS+AHGIIRS VL++ L  +F+GN+VL  VL RKP
Sbjct: 105 MPPSCTNSTQENNGSRVCLPLSKMPISVAHGIIRSVVLLVILGVAFLGNVVLGYVLHRKP 164

Query: 61  QLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVN 120
            LLQVTNRFIFNLLVTDLLQ++LVAPWVV+T++P FWPLN HFCTALVSLTHLFAFASVN
Sbjct: 165 NLLQVTNRFIFNLLVTDLLQVALVAPWVVSTAIPFFWPLNIHFCTALVSLTHLFAFASVN 224

Query: 121 TIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNA 180
           TIVVVSVDRYL+IIHPLSYPSKMT RR Y+LLYGTWI A LQSTPPLYGWG A FD+RNA
Sbjct: 225 TIVVVSVDRYLTIIHPLSYPSKMTNRRSYILLYGTWIAAFLQSTPPLYGWGHATFDDRNA 284

Query: 181 LCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKD 240
            CSMIWGASP+YT++SVVSF+VIPL VMIACYSVVF AARRQ ALLY  K H LEVRV+D
Sbjct: 285 FCSMIWGASPAYTVVSVVSFLVIPLGVMIACYSVVFGAARRQQALLYKAKSHRLEVRVED 344

Query: 241 CV 242
            V
Sbjct: 345 SV 346
```

```
 Score =  195 bits (496), Expect = 1e-48
 Identities = 93/143 (65%), Positives = 104/143 (72%), Gaps = 4/143 (2%)

Query: 260 MNIPEXXXXXXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET 319
           M IPE                 CY+CKAA+VIF+II +YVLSLGPYCFLAVLAVWVD++T
Sbjct: 473 MRIPESSPPSRRNSTSDPPLPPCYECKAARVIFVIISTYVLSLGPYCFLAVLAVWVDIDT 532

Query: 320 QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDMLXXXXXXXXXXXXDSHPDL-- 377
           +VPQWVITIIIWLFFLQCCIHPYVYGYMHK+IKKEIQ++L            DSHPDL
Sbjct: 533 RVPQWVITIIIWLFFLQCCIHPYVYGYMHKSIKKEIQEVLKKLICKKSPPVEDSHPDLHE 592

Query: 378 --PGTEGGTEGKIVPSYDSATFP 398
              GTEGG EGK VPS+DSAT P
Sbjct: 593 TEAGTEGGIEGKAVPSHDSATSP 615
```

>gi|34881783|ref|XP_229186.2| similar to G protein-coupled receptor 101 [Rattus
Length = 508

```
 Score =  404 bits (1037), Expect = e-111
 Identities = 191/242 (78%), Positives = 212/242 (87%)

Query: 1   MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKP 60
           M S+CTNST+E+NSS  C+PLSKMPIS+AHGIIRS VL+I L  +FVGN+VL  VL RKP
Sbjct: 1   MPSSCTNSTQENNSSRVCLPLSKMPISIAHGIIRSVVLLIILGVAFVGNVVLGYVLHRKP 60
```

```
Query: 61  QLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVN 120
           LLQVTNRFIFNLLVTDLLQ++LVAPWVV+T++P FWPLN HFCTALVSLTHLFAFASVN
Sbjct: 61  HLLQVTNRFIFNLLVTDLLQVALVAPWVVSTAIPFFWPLNIHFCTALVSLTHLFAFASVN 120

Query: 121 TIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNA 180
           TIVVVS+DRYLSIIHPLSYPSKMT RR Y+LLYGTWI A LQSTPPLYGWG A FD+RNA
Sbjct: 121 TIVVVSIDRYLSIIHPLSYPSKMTNRRSYILLYGTWIAAFLQSTPPLYGWGHATFDDRNA 180

Query: 181 LCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKD 240
           CSMIWG SP+YT++SVVSF+VIPL VMIACYSVVF AARRQ ALLY  K H  +VRVKD
Sbjct: 181 FCSMIWGDSPAYTVVSVVSFLVIPLGVMIACYSVVFGAARRQQALLYKAKSHRFQVRVKD 240

Query: 241 CV 242
           V
Sbjct: 241 SV 242


 Score =  197 bits (501), Expect = 3e-49
 Identities = 93/143 (65%), Positives = 106/143 (74%), Gaps = 4/143 (2%)

Query: 260 MNIPEXXXXXXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVET 319
           M IPE                CY+CKAA+VIFIIIFSYVLSLGPYCFLAVLAVWVD+++
Sbjct: 366 MRIPESRPPSRRNSTSNPPLPPCYECKAARVIFIIIFSYVLSLGPYCFLAVLAVWVDIDS 425

Query: 320 QVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDMLXXXXXXXXXXXXDSHPDL-- 377
           QVPQWVITIIIWLFFLQCC+HPYVYGYMHK+IKKEI+++L            DSHP+L
Sbjct: 426 QVPQWVITIIIWLFFLQCCVHPYVYGYMHKSIKKEIKEVLKKLTCKKSTSVDDSHPELRE 485

Query: 378 --PGTEGGTEGKIVPSYDSATFP 398
             GTEGGTEGK +PS+DSAT P
Sbjct: 486 TEAGTEGGTEGKAIPSHDSATSP 508


>gi|24476016|ref|NP_722561.1|  G protein-coupled receptor 161; G-protein coupled
          sapiens]
 gi|20381353|gb|AAH28163.1|  G protein-coupled receptor 161 [Homo sapiens]
 gi|30526190|gb|AAP32300.1|  G-protein coupled receptor RE2 [Homo sapiens]
          Length = 529

 Score =  137 bits (346), Expect = 3e-31
 Identities = 87/323 (26%), Positives = 151/323 (46%), Gaps = 21/323 (6%)

Query: 37  VLVIFLAASFVGNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLF 96
           V+ IF+    +GN+V+ + L +K  LL ++N+F+F+L +++ L    LV P+VV +S+
Sbjct: 36  VITIFVC---LGNLVIVVTLYKKSYLLTLSNKFVFSLTLSNFLLSVLVLPFVVTSSIRRE 92

Query: 97  WPLNSHFCTALVSLTHLFAFASVNTIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTW 156
           W     +C     L  L + AS+ T+ V+++DRY ++++P+ YP K+T  R  + L   W
Sbjct: 93  WIFGVVWCNFSALLYLLISSASMLTLGVIAIDRYYAVLYPMVYPMKITGNRAVMALVYIW 152

Query: 157 IVAILQSTPPLYGWGQAAFDERNALCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVF 216
           + +++   PPL+GW     FDE   +C   W   P YT     +  + P +VM+ CY  +F
Sbjct: 153 LHSLIGCLPPLFGWSSVEFDEFKWMCVAAWHREPGYTAFWQIWCALFPFLVMLVCYGFIF 212

Query: 217 CAARRQHALLYNVKRHSLEVRVKDCVXXXXXXXXXXXXXXXXXMNIPEXXXXXXXXXXXX 276
             AR               V+ +                     +
```

```
Sbjct: 213 RVAR---------------VKARKVHCGTVVIVEEDAQRTGRKNSSTSTSSSGSRRNAFQ 257

Query: 277 XXXXXCYQCKAAKVIFIIIFSYVLSLGPY-CFLAVLAVWVDVETQVPQWVITIIIWLFFL 335
                QCKA   I +++ +++++ GPY   +A  A+W   ++ V   + T   WL F
Sbjct: 258 GVVYSANQCKALITILVVLGAFMVTWGPYMVVIASEALW--GKSSVSPSLETWATWLSFA 315

Query: 336 QCCIHPYVYGYMHKTIKKEIQDM 358
             HP +YG  +KT++KE+  M
Sbjct: 316 SAVCHPLIYGLWNKTVRKELLGM 338
```

>gi|6677701|ref|NP_031395.1| G protein-coupled receptor 161; G-protein coupled r
sapiens]
gi|3659903|gb|AAC61598.1| G-protein coupled receptor RE2 [Homo sapiens]
Length = 407

```
 Score =  134 bits (336), Expect = 4e-30
 Identities = 88/330 (26%), Positives = 153/330 (46%), Gaps = 20/330 (6%)

Query: 32  IIRSTVLVIFLAASFV--GNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISLVAPWVV 89
           +I +  + I +   FV  GN+V+ + L +K  LL ++N+F+F+L +++ L   LV P+VV
Sbjct: 26  VIITQFIAIIVITIFVCLGNLVIVVTLYKKSYLLTLSNKFVFSLTLSNFLLSVLVLPFVV 85

Query: 90  ATSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVSVDRYLSIIHPLSYPSKMTQRRGY 149
            +S+   W     +C     L  L + AS+ T+ V+++DRY ++++P+ YP K+T  R
Sbjct: 86  TSSIRREWIFGVVWCNFSALLYLLISSASMLTLGVIAIDRYYAVLYPMVYPMKITGNRAV 145

Query: 150 LLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIWGASPSYTILSVVSFIVIPLIVMI 209
           + L   W+ +++   PPL+GW    FDE   +C   W   P YT    +  + P +VM+
Sbjct: 146 MALVYIWLHSLIGCLPPLFGWSSVEFDEFKWMCVAAWHREPGYTAFWQIWCALFPFLVML 205

Query: 210 ACYSVVFCAARRQHALLYNVKRHSLEVRVKDCVXXXXXXXXXXXXXXXXXMNIPEXXXXX 269
            CY  +F  AR               V+ +                     +
Sbjct: 206 VCYGFIFRVAR---------------VKARKVHCGTVVIVEEDAQRTGRKNSSTSTSSSG 250

Query: 270 XXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPY-CFLAVLAVWVDVETQVPQWVITI 328
                       QCKA   I +++ +++++ GPY   +A  A+W   ++ V   + T
Sbjct: 251 SRRNAFQGVVYSANQCKALITILVVLGAFMVTWGPYMVVIASEALW--GKSSVSPSLETW 308

Query: 329 IIWLFFLQCCIHPYVYGYMHKTIKKEIQDM 358
             WL F     HP +YG  +KT++KE+  M
Sbjct: 309 ATWLSFASAVCHPLIYGLWNKTVRKELLGM 338
```

>gi|29611554|gb|AAO85088.1| G protein-coupled receptor GPR101 [Mus musculus]
Length = 77

```
 Score =  129 bits (324), Expect = 9e-29
 Identities = 62/76 (81%), Positives = 68/76 (89%)

Query: 141 SKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIWGASPSYTILSVVSF 200
           SKMT RR Y+LLYGTWI A LQSTPPLYGWG A FD+RNA CSMIWGASP+YT++SVVSF
Sbjct: 1   SKMTNRRSYILLYGTWIAAFLQSTPPLYGWGHATFDDRNAFCSMIWGASPAYTVVSVVSF 60

Query: 201 IVIPLIVMIACYSVVF 216
           +VIPL VMIACYSVVF
Sbjct: 61  LVIPLGVMIACYSVVF 76
```

>gi|8843925|gb|AAF80168.1| alpha 1a-adrenoceptor isoform 2 [Oryctolagus cuniculus]
Length = 429

Score = 124 bits (311), Expect = 3e-27
Identities = 90/353 (25%), Positives = 160/353 (45%), Gaps = 18/353 (5%)

```
Query: 7   NSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKPQLLQVT 66
            N++  SN +H         P++++  I+   +L   +    +GNI++ L +     L  VT
Sbjct: 7   NASDSSNCTH-----PPAPVNISKAILLGVILGGLILFGVLGNILVILSVACHRHLHSVT 61

Query: 67  NRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVS 126
            + +I NL V DLL  S V P+     +  +W     FC    ++  L   AS+ ++ V+S
Sbjct: 62  HYYIVNLAVADLLLTSTVLPFSAIFEILGYWAFGRVFCNIWAAVDVLCCTASIISLCVIS 121

Query: 127 VDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIW 186
            +DRY+ + +PL YP+ +TQRRG   L   W  +++ S  PL+GW Q A D+   +C +
Sbjct: 122 IDRYIGVSYPLRYPTIVTQRRGLRALLCVWAFSLVISVGPLFGWRQPAPDDE-TICQI-- 178

Query: 187 GASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKDCVXXXX 246
               P Y + S +    +PL +++A Y  V+  A+R+   L    +  L+    D
Sbjct: 179 NEEPGYVLFSALGSFYVPLTIILAMYCRVYVVAKRESRGL----KSGLKTDKSDSEQVTL 234

Query: 247 XXXXXXXXXXXXXMNIPEXXXXXXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPYC 306
                       +   +                  + KAAK + I++  +VL   P+
Sbjct: 235 RIHRKNAPAGGSGVASAKNKTHFSVRLLKFSR-----EKKAAKTLGIVVGCFVLCWLPFF 289

Query: 307 FLAVLAVWVDVETQVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDML 359
            +  +  + + P+ V  I+ WL +L  CI+P +Y    +  KK  Q++L
Sbjct: 290 LVMPIGSFFP-DFKPPETVFKIVFWLGYLNSCINPIIYPCSSQEFKKAFQNVL 341
```

>gi|3023219|sp|O02824|A1AA_RABIT Alpha-1A adrenergic receptor (Alpha 1A-adrenocept adrenergic receptor)
gi|2198745|gb|AAB61334.1| alpha 1a-adrenoceptor [Oryctolagus cuniculus]
Length = 466

Score = 124 bits (311), Expect = 3e-27
Identities = 90/353 (25%), Positives = 160/353 (45%), Gaps = 18/353 (5%)

```
Query: 7   NSTRESNSSHTCMPLSKMPISLAHGIIRSTVLVIFLAASFVGNIVLALVLQRKPQLLQVT 66
            N++  SN +H         P++++  I+   +L   +    +GNI++ L +     L  VT
Sbjct: 7   NASDSSNCTH-----PPAPVNISKAILLGVILGGLILFGVLGNILVILSVACHRHLHSVT 61

Query: 67  NRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVS 126
            + +I NL V DLL  S V P+     +  +W     FC    ++  L   AS+ ++ V+S
Sbjct: 62  HYYIVNLAVADLLLTSTVLPFSAIFEILGYWAFGRVFCNIWAAVDVLCCTASIISLCVIS 121

Query: 127 VDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYGWGQAAFDERNALCSMIW 186
            +DRY+ + +PL YP+ +TQRRG   L   W  +++ S  PL+GW Q A D+   +C +
Sbjct: 122 IDRYIGVSYPLRYPTIVTQRRGLRALLCVWAFSLVISVGPLFGWRQPAPDDE-TICQI-- 178

Query: 187 GASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALLYNVKRHSLEVRVKDCVXXXX 246
               P Y + S +    +PL +++A Y  V+  A+R+   L    +  L+    D
Sbjct: 179 NEEPGYVLFSALGSFYVPLTIILAMYCRVYVVAKRESRGL----KSGLKTDKSDSEQVTL 234

Query: 247 XXXXXXXXXXXXXMNIPEXXXXXXXXXXXXXXXXCYQCKAAKVIFIIIFSYVLSLGPYC 306
```

```
                          +   +                      + KAAK + I++  +VL   P+
Sbjct: 235 RIHRKNAPAGGSGVASAKNKTHFSVRLLKFSR-----EKKAAKTLGIVVGCFVLCWLPFF 289

Query: 307 FLAVLAVWVDVETQVPQWVITIIIWLFFLQCCIHPYVYGYMHKTIKKEIQDML 359
            +  +  +   + + P+ V  I+ WL +L  CI+P +Y    +  KK  Q++L
Sbjct: 290 LVMPIGSFFP-DFKPPETVFKIVFWLGYLNSCINPIIYPCSSQEFKKAFQNVL 341


>gi|26788019|emb|CAC94897.2|  SI:bZ20I5.4 (novel protein similar to human G-protein
          receptor RE2) [Danio rerio]
         Length = 526

 Score =  124 bits (311), Expect = 3e-27
 Identities = 86/312 (27%), Positives = 147/312 (47%), Gaps = 18/312 (5%)

Query: 47  VGNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISLVAPWVVATSVPLFWPLNSHFCTA 106
           +GN+V+ + L +KP LL  +N+F+F+L  ++LL   L+ P+VVA+SV   W     +C
Sbjct: 39  LGNLVIVVTLYKKPYLLTPSNKFVFSLTSSNLLLSVLMLPFVVASSVRRDWMFGVVWCNF 98

Query: 107 LVSLTHLFAFASVNTIVVVSVDRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPP 166
              L  L + +S+ T+  +++DRY ++++P+ YP K+T  R  L +   W+ +++   PP
Sbjct: 99  TALLHLLVSSSSMLTLGAIAIDRYYAVLYPMIYPMKITGNRAVLAIVYIWLHSLVGCLPP 158

Query: 167 LYGWGQAAFDERNALCSMIWGASPSYTILSVVSFIVIPLIVMIACYSVVFCAARRQHALL 226
           L+GW     FD      C++ W    SYT   V    ++PL+ M+ CY V+F  AR +   +
Sbjct: 159 LFGWSSFEFDRFKWTCTVSWHKEISYTAFWVTWCCLLPLVAMLVCYGVIFRVARIKARKV 218

Query: 227 YNVKRHSLEVRVKDCVXXXXXXXXXXXXXXXXXMNIPEXXXXXXXXXXXXXXXXXCYQCK 286
           Y              C                          +                QCK
Sbjct: 219 Y-------------CGSVVVSQEESSSQNNGRKNSNTSTSSSGSRKSLIYSGS----QCK 261

Query: 287 AAKVIFIIIFSYVLSLGPYCFLAVLAVWVDVETQVPQWVITIIIWLFFLQCCIHPYVYGY 346
           A   I +++ +++ + GPY  +      +   +  PQ V T++ WL F      HP +YG
Sbjct: 262 AFITILVVLGTFLTTWGPYVVVISTEALLGKNSVSPQ-VETLVSWLSFTSAVCHPLIYGL 320

Query: 347 MHKTIKKEIQDM 358
            +KT++KE+  M
Sbjct: 321 WNKTVRKELLGM 332
```

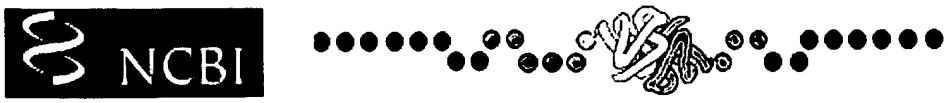
```
  Database: All non-redundant GenBank CDS
  translations+PDB+SwissProt+PIR+PRF
    Posted date:  Jan 5, 2004 12:06 AM
  Number of letters in database: 521,258,069
  Number of sequences in database:  1,591,231

Lambda     K      H
   0.328    0.139    0.442

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
```

Number of Hits to DB: 33,689,817
Number of Sequences: 1591231
Number of extensions: 1154639
Number of successful extensions: 5029
Number of sequences better than 10.0: 734
Number of HSP's better than 10.0 without gapping: 402
Number of HSP's successfully gapped in prelim test: 332
Number of HSP's that attempted gapping in prelim test: 3938
Number of HSP's gapped (non-prelim): 964
length of query: 398
length of database: 521,258,069
effective HSP length: 127
effective length of query: 271
effective length of database: 319,171,732
effective search space: 86495539372
effective search space used: 86495539372
T: 11
A: 40
X1: 15 (7.1 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 40 (21.8 bits)
S2: 74 (33.1 bits)



1: NP_473362. G protein-coupled...[gi:16876435]

BLink, Domains, Links

```
LOCUS       NP_473362                 508 aa            linear   PRI 05-OCT-2003
DEFINITION  G protein-coupled receptor 101 [Homo sapiens].
ACCESSION   NP_473362
VERSION     NP_473362.1  GI:16876435
DBSOURCE    REFSEQ: accession NM_054021.1
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 508)
  AUTHORS   Takeda,S., Kadowaki,S., Haga,T., Takaesu,H. and Mitaku,S.
  TITLE     Identification of G protein-coupled receptor genes from the human
            genome sequence
  JOURNAL   FEBS Lett. 520 (1-3), 97-101 (2002)
  MEDLINE   22040266
   PUBMED   12044878
REFERENCE   2  (residues 1 to 508)
  AUTHORS   Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,
            Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.
  TITLE     Discovery and mapping of ten novel G protein-coupled receptor genes
  JOURNAL   Gene 275 (1), 83-91 (2001)
  MEDLINE   21458557
   PUBMED   11574155
COMMENT     PROVISIONAL REFSEQ: This record has not yet been subject to final
            NCBI review. The reference sequence was derived from AF411115.1.

            Summary: G protein-coupled receptors (GPCRs, or GPRs) contain 7
            transmembrane domains and transduce extracellular signals through
            heterotrimeric G proteins.[supplied by OMIM].
FEATURES             Location/Qualifiers
     source          1..508
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq25-27.1"
     Protein         1..508
                     /product="G protein-coupled receptor 101"
     Region          48..240
                     /region_name="7 transmembrane receptor (rhodopsin family)"
                     /note="7tm_1"
                     /db_xref="CDD:pfam00001"
     variation       124
                     /allele="V"
                     /allele="L"
                     /db_xref="dbSNP:1190736"
     variation       376
```

```
                     /allele="P"
                     /allele="L"
                     /db_xref="dbSNP:5931046"
     Region          396..454
                     /region_name="7 transmembrane receptor (rhodopsin family)"
                     /note="7tm_1"
                     /db_xref="CDD:pfam00001"
     CDS             1..508
                     /gene="GPR101"
                     /coded_by="NM_054021.1:1..1527"
                     /note="go_component: integral to membrane [goid 0016021]
                     [evidence IEA];
                     go_function: receptor activity [goid 0004872] [evidence
                     IEA];
                     go_function: rhodopsin-like receptor activity [goid
                     0001584] [evidence IEA];
                     go_process: G-protein coupled receptor protein signaling
                     pathway [goid 0007186] [evidence IEA]"
                     /db_xref="GeneID:83550"
                     /db_xref="LocusID:83550"
                     /db_xref="MIM:300393"
ORIGIN
        1 mtstctnstr esnsshtcmp lskmpislah giirstvlvi flaasfvgni vlalvlqrkp
       61 qllqvtnrfi fnllvtdllq islvapwvva tsvplfwpln shfctalvsl thlfafasvn
      121 tivvvsvdry lsiihplsyp skmtqrrgyl llygtwivai lqstpplygw gqaafderna
      181 lcsmiwgasp sytilsvvsf iviplivmia cysvvfcaar rqhallynvk rhslevrvkd
      241 cvenedeega ekkeefqdes efrrqhegev kakegrmeak dgslkakegs tgtsessvea
      301 rgseevress tvasdgsmeg kegstkveen smkadkgrte vnqcsidlge ddmefgeddi
      361 nfseddveav nipeslppsr rnsnsnpplp rcyqckaakv ifiiifsyvl slgpycflav
      421 lavwvdvetq vpqwvitiii wlfflqccih pyvygymhkt ikkeiqdmlk kffckekppk
      481 edshpdlpgt eggtegkivp sydsatfp
//
```


Nov 21 2003 07:32:41

Research

Phylogenetic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands

Patrick Joost and Axel Methner

Address: Research Group Protective Signaling, Zentrum für Molekulare Neurobiologie Hamburg and Klinik für Neurologie, Universitätskrankenhaus Eppendorf, Martinistrasse 52, D-20251 Hamburg, Germany.

Correspondence: Axel Methner. E-mail: methner@uke.uni-hamburg.de

Published: 17 October 2002

Genome Biology 2002, 3(11):research0063.1–0063.16

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2002/3/11/research/0063

(Print ISSN 1465-6906; Online ISSN 1465-6914)

Received: 19 June 2002
Revised: 7 August 2002
Accepted: 18 September 2002

Abstract

Background: G-protein-coupled receptors (GPCRs) are the largest and most diverse family of transmembrane receptors. They respond to a wide range of stimuli, including small peptides, lipid analogs, amino-acid derivatives, and sensory stimuli such as light, taste and odor, and transmit signals to the interior of the cell through interaction with heterotrimeric G proteins. A large number of putative GPCRs have no identified natural ligand. We hypothesized that a more complete knowledge of the phylogenetic relationship of these orphan receptors to receptors with known ligands could facilitate ligand identification, as related receptors often have ligands with similar structural features.

Results: A database search excluding olfactory and gustatory receptors was used to compile a list of accession numbers and synonyms of 81 orphan and 196 human GPCRs with known ligands. Of these, 241 sequences belonging to the rhodopsin receptor-like family A were aligned and a tentative phylogenetic tree constructed by neighbor joining. This tree and local alignment tools were used to define 19 subgroups of family A small enough for more accurate maximum-likelihood analyses. The secretin receptor-like family B and metabotropic glutamate receptor-like family C were directly subjected to these methods.

Conclusions: Our trees show the overall relationship of 277 GPCRs with emphasis on orphan receptors. Support values are given for each branch. This approach may prove valuable for identification of the natural ligands of orphan receptors as their relation to receptors with known ligands becomes more evident.

Background

G-protein-coupled receptors (GPCRs) are the largest and most diverse family of transmembrane receptors. They respond to a wide range of stimuli including small peptides, lipid analogs, amino-acid derivatives, and sensory stimuli such as light, taste and odor [1], and transmit signals to the interior of the cell through interaction with heterotrimeric G proteins. Certain amino-acid residues of this receptor family are well conserved and approaches exploiting this, such as low-stringency hybridization and degenerate PCR, have been used to clone new members of this large superfamily [2-4]. Many of these putative receptors share GPCR structural motifs, but still lack a defined physiologically relevant ligand. One strategy to identify the natural ligand of these so-called

orphan receptors uses changes in second-messenger activation in cells stably expressing the receptor in response to tissue extracts expected to contain the natural ligand [5]. In a second step, these extracts are tested and fractionated to purity, before being analyzed by mass spectrometry. This strategy led to the identification of several novel bioactive peptides or peptide families (for review see [6]). The identification of these natural ligands is likely to give further insight into the physiological role of these receptors and advance the design of pharmacologically active receptor agonists or antagonists. This is of particular interest, as GPCRs are the most targeted protein superfamily in pharmaceutical research [7]. Better prediction of the presumed chemical class or structure of the ligand facilitates the identification of orphan receptors by the strategy described above, as the ligand purification process can be tailored more specifically to the assumed class of substances.

Phylogenetic analysis of receptor relationships has already been used to elucidate the chemical nature of receptor ligands. The identification of sphingosine 1-phosphate as the ligand for the GPCR EDG-1 led to the prediction that EDG-3, EDG-5, EDG-6 and EDG-8 have the same ligand [8-11]. In contrast, phylogenetically distinct members of the EDG cluster - EDG-2, EDG-4 and EDG-7 - are receptors for the similar but distinct ligand lysophosphatidic acid (LPA) [12-14]. Neuromedin U, a potent neuropeptide that causes contraction of smooth muscle, was correctly predicted phylogenetically to be the ligand of the orphan GPCR FM3 (NMUR) [15]. Not only the ligand, but also the pharmacology of a novel receptor for histamine, was predicted and confirmed through phylogeny [16]. GPR86, related to the ADP receptor P2Y12, was similarly recently shown to bind ADP [17], and UDP-glucose, a molecule involved in carbohydrate biosynthesis, was shown to be the ligand for the related receptor KIAA0001 [18].

Mammalian GPCRs were previously classified by phylogeny into three families [19,20]: the rhodopsin receptor-like family (A), the secretin receptor-like receptor family (B) and the metabotropic glutamate receptor family (C). These results were generated by neighbor joining, a fast distance-based method suited for large datasets, but influenced by methodological flaws that can in part be overcome by methods not generally applied previously.

In this work, we compiled an exhaustive list that includes all available synonyms and accession numbers of 196 human GPCRs with known ligands and 84 human orphan receptors. The 241 sequences belonging to family A were aligned, and a tentative tree constructed by neighbor joining with 1,000 bootstrap steps. Subgroups of family A defined by this tree and sequences from families B and C were then used for more accurate phylogenetic analysis by state-of-the-art techniques. From this analysis, we tried to predict possible ligands for orphan receptors.

Results and discussion

We set out to define the phylogenetic relationship of human GPCRs by state-of-the-art tools, assuming that the identification of cognate ligands of orphan receptors will be facilitated by a more complete knowledge of their relationship within the large and diverse superfamily.

Database mining and multiple sequence alignment

Most receptors were identified by different groups; therefore, many confusing names and synonyms exist. We adhered to SWISS-PROT names where possible, and compiled a list including all available synonyms and accession numbers of 196 human GPCRs with known ligands and 84 human orphan receptors (Table 1 shows all receptors mentioned in this work; the complete list is supplied as an additional data file with the online version of this paper). Gustatory and olfactory receptors were omitted. Multiple protein sequences were aligned and the extremely variable amino termini upstream of the first transmembrane domain and carboxyl termini downstream of the seventh transmembrane domain were deleted to avoid length heterogeneity (see Figure 1). The deleted regions contained no significant sequence conservation.

Phylogenetic analysis

Because of the large number of sequences in family A, we had to use a combination of computational methods to accomplish the best possible description of their phylogenetic relationship. In a first step we used the distance-based neighbor-joining method as the only one computationally feasible. Neighbor joining has been shown to be efficient at recovering the correct tree topology [21], but is greatly influenced by methodological errors, for example, the sampling error [22]. This can in part be overcome by bootstrapping, a method of testing the reliability of a dataset by the creation of pseudoreplicate datasets by resampling. Bootstrapping assesses whether stochastic effects have influenced the distribution of amino acids [23]. In previous publications on this topic, bootstrapping has not been generally used.

We generated a neighbor-joining tree of family-A sequences, and considered tree branches to be confirmed if they were found in more than 500 of 1,000 bootstrap steps (Figure 2). The same branching pattern was found by least squares (data not shown) as implemented in FITCH [24], but it was not possible to compute enough bootstrap steps with the equipment used. The remaining sequences of unconfirmed branches were then assigned to existing branches according to results obtained with the local alignment tool BLASTP (see Additional data files) [25] to account for similarities in parts of the sequences not sufficient for repeated global alignment. The *p*-value was used as a measure of similarity.

As this strategy still left four subgroups too large for detailed analyses, we recalculated neighbor-joining trees and in some cases least-square trees of these sequences to create

Table 1

List of example receptor names, accession numbers and abbreviations

Receptor	Group	Accession no.	Names and synonyms
Human GPCR - Family A			
ADMR	**A02**	**O15218**	**Adrenomedullin receptor, Am-R**
APJ	A03	P35414	Apelin receptor, Apj, Agtrl1
CML1	**A08**	**Q99788**	**Chemokine receptor-like 1, Dez, Chemr23, Ch23, Cmklr1**
CML2	**A02**	**Q99527**	**Chemokine receptor-like 2, flow-induced endothelial G protein-coupled receptor, Feg-1, Gpr30, Cmkrl2, Dry12, Cepr**
DUFF	A02	Q16570	Duffy antigen, Fy glycoprotein, glycoprotein D, Gpfy, Fy, Gpd, Darc
EDG1	A13	P21453	Endothelial differentiation, Sphingosine 1-phosphate receptor, Lp-B1
EDG2	**A13**	**Q92633**	**Endothelial differentiation, lysophosphatidic acid receptor, Lp-A1, Vzg-1**
EDG3	**A13**	**Q99500**	**Endothelial differentiation, lysosphingolipid receptor, Lp-B3**
EDG4	A13	NM_004720	Endothelial differentiation, lysophosphatidic acid receptor, Lp-A2
EDG5	A13	NP_004221	Endothelial differentiation, sphingolipid receptor, Lp-B2, H218, Agr16
EDG6	A13	AJ000479	Endothelial differentiation, lysosphingolipid receptor, Lp-C1
EDG7	A13	NP_036284	Endothelial differentiation, lysophosphatidic acid receptor, Lp-A3
EDG8	A13	NP_110387	Endothelial differentiation, sphingosine 1-phosphate receptor, Lp-B4
ETBR-LP2	**A07**	**Y16280**	**Endothelin B receptor-like protein-2, Etbrlp2, Ebp2, Cns2**
FSHR	A10	P23945	Follicle stimulating hormone receptor, Fsh-R, follitropin receptor
GPR	**A06**	**NM_007223**	**G protein-coupled receptor**
GPR1	**A08**	**P46091**	**G protein-coupled receptor Gpr1**
GPR3	**A13**	**P46089**	**G protein-coupled receptor, Acca orphan receptor**
GPR6	**A13**	**P46095**	**G protein-coupled receptor 6**
GPR7	**A04**	**P48145**	**G protein-coupled receptor 7**
GPR8	**A04**	**P48146**	**G protein-coupled receptor 8**
GPR25	**A03**	**NM_005298**	**G protein-coupled receptor 25**
GPR27	**A18**	**NM_018971**	**G protein-coupled receptor 27, Sreb1**
GPR34	**A12**	**NM_005300**	**G protein-coupled receptor, Gpry**
GPR35	**A15**	**NM_005301**	**G protein-coupled receptor 35**
GPR37	**A07**	**NM_005302**	**G protein-coupled receptor 37, Endothelin receptor type B-like, Cns1**
GPR39	**A07**	**O43194**	**G protein-coupled receptor Gpr39**
GPR40	**A11**	**O14842**	**G protein-coupled receptor Gpr40**
GPR41	**A11**	**O14843**	**G protein-coupled receptor Gpr41, Hia-R**
GPR42	**A11**	**O15529**	**G protein-coupled receptor Gpr42**
GPR43	**A11**	**O15552**	**G protein-coupled receptor Gpr43**
GPR44	**A08**	**AAD21055**	**G protein-coupled receptor 44**
GPR44	**A08**	**AAD21055**	**G protein-coupled receptor 44**
GPR48	A10	NM_018490	G protein-coupled receptor 48
GPR49	A10	NM_003667	G protein-coupled receptor 49, Hg38, G protein-coupled receptor 67, Fex
GPR52	**A18**	**Q9Y2T5**	**G protein-coupled receptor Gpr52**
GPR55	**A15**	**NM_005683**	**G protein-coupled receptor 55**
GPR57	**A17**	**NM_014627**	**G protein-coupled receptor 57**
GPR58	**A17**	**NM_014626**	**G protein-coupled receptor 58**
GPR61	**A18**	**AF317652**	**G protein-coupled receptor 61**
GPR62	**A18**	**AF317653**	**G protein-coupled receptor 62**
GPR63	**A18**	**AF317654**	**G protein-coupled receptor 63**
GPR72	**A09**	**NM_016540**	**G protein-coupled receptor 72, Jp05**
GPR73	**A09**	**AAE24084**	**G protein-coupled receptor 73**
GPR75	**A09**	**NM_006794**	**G protein-coupled receptor 75**
GPR80	**A11**	**AF411109**	**G protein-coupled receptor 80**
GPR81	**A11**	**AF411110**	**G protein-coupled receptor 81**
GPR85	**A18**	**NM_018970**	**G protein-coupled receptor 85, Sreb2**
GPR86	A12	NP_076403	Adp receptor
GPR87	**A12**	**NM_023915**	**G protein-coupled receptor 87**
GPR88	**A18**	**NM_022049**	**G protein-coupled receptor 88**
GPR91	**A11**	**NM_033050**	**G protein-coupled receptor 91**

Table 1 *(continued)*

Receptor	Group	Accession No.	Names & Synonyms
GPR101	**A18**	**NM_054021**	**G protein-coupled receptor 101**
GPR102	**A17**	**NM_053278**	**G protein-coupled receptor 102**
GPR103	**A06**	**AF411117**	**G protein-coupled receptor 103**
GPRC	A13	P47775	Gpr12
GPRF	**A03**	**P49685**	**Gpr15, Bob**
GPRJ	**A09**	**Q15760**	**Gpr19, Gpr-Nga**
GPRL	**A18**	**Q99679**	**Gpr21**
GPRM	**A06**	**Q99680**	**Gpr22**
GPRV	**A11**	**O00270**	**Gpr31**
GPRW	**A08**	**O75388**	**Gpr32**
HM74	**A11**	**P49019**	**G protein-coupled receptor Hm74**
KI01	A12	Q15391	Udp-Glucose receptor, Kiaa0001
LSHR	A10	P22888	Lutropin-choriogonadotropic hormone receptor, Lh/Cg-R, Lsh-R, luteinizing hormone receptor, Lhcgr, Lhrhr, Lcgr
MAS	**A08**	**P04201**	**Mas proto-oncogene, Mas1**
ML1A	A09	P48039	Melatonin receptor Type 1a, Mel-1a-R, Mtnr1a
ML1B	A09	P49286	Melatonin receptor Type 1b, Mel-1b-R, Mtnr1b
ML1X	**A09**	**Q13585**	**Melatonin-related receptor, H9, Gpr50**
MRG	**A08**	**P35410**	**Mas-related G protein-coupled receptor**
NMU1R	A07	AF272362	Neuromedin U receptor 1, Nmur1, Gpr66, Fm-3
NTR1	A07	P30989	Neurotensin receptor Type 1, Nt-R-1, Ntsr1, Ntrr
NTR2	A07	O95665	Neurotensin receptor Type 2, Nt-R-2, levocabastine-sensitive neurotensin receptor, Ntr2 receptor, Ntsr2
NY1R	A09	P25929	Neuropeptide Y receptor Type 1, Npy1-R, Npy1r, Npyr, Npyy1
NY2R	A09	P49146	Neuropeptide Y receptor Type 2, Npy2-R, Npy2r
NY4R Npy4r	A09	P50391	Neuropeptide Y receptor Type 4, Npy4-R, Pancreatic Polypeptide receptor 1, Pp1, Ppyr1,
P2Y5	**A15**	**P43657**	**P2y purinoceptor 5, P2y5, purinergic receptor 5, P2ry5, 6h1**
P2Y7	A05	Q15722	P2y purinoceptor 7, P2y7, Leukotriene B4 receptor, Chemoattractant receptor-like 1, P2ry7, P2y7, Gpr16, Cmkrl1, Ltb4r
P2Y9	**A15**	**Q99677**	**P2y purinoceptor 9, P2y9, purinergic receptor 9, Gpr23, P2ry9**
P2Y10	**A15**	**AF000545**	**Putative purinergic receptor P2y10**
P2Y12	A12	AF313449	Adp receptor, Sp1999
PAFR	A12	P25105	Platelet Activating Factor receptor, Paf-R, Ptafr
PNR	**A17**	**AF021818**	**Putative neurotransmitter receptor**
PSP24	**A18**	**U92642**	**High-affinity lysophosphatidic acid receptor homolog, Gpr45**
RDC1	**A02**	**P25106**	**G protein-coupled receptor Rdc1 homolog**
RE2	**A18**	**AF091890**	**G protein-coupled receptor Re2**
SALPR	**A05**	**NM_016568**	**Somatostatin and angiotensin-like peptide receptor, Loc51289**
SREB3	**A18**	**NM_018969**	**Super conserved receptor expressed in brain 3**
TM7SF1	**A01**	**AF027826**	**Putative seven pass transmembrane protein**
TSHR	A10	P16473	Thyroid stimulating hormone receptor, thyrotropin receptor , Tsh-R
Human GPCR - Family B			
EMR1	**B**	**Q14246**	**Cell surface glycoprotein emr1, Emr1 hormone receptor**
EMR2	**B**	**AF114491**	**Egf-like module Emr2**
EMR3	**B**	**AF239764**	**Egf-like module-containing mucin-like receptor Emr3**
BAI1	**B**	**O14514**	**Brain-specific angiogenesis inhibitor 1**
BAI2	**B**	**O60241**	**Brain-specific angiogenesis inhibitor 2**
BAI3	**B**	**O60242**	**Brain-specific angiogenesis inhibitor 3, Kiaa0550**
GPR56	**B**	**NM_005682**	**G protein-coupled receptor 56**
Human GPCR - Family C			
GPRC5B	**C**	**NM_016235**	**G PROTEIN-COUPLED RECEPTOR, FAMILY C, GROUP 5, MEMBER B, GPRC5B**
GPRC5C	**C**	**NM_018653**	**G protein-coupled receptor, family C, group 5, member C, Gprc5c**
GPRC5D	**C**	**NM_018654**	**G protein-coupled receptor, family C, group 5, member D, Gprc5d**

A complete list is supplied as additional data file. Orphan receptors are shown in bold.

subgroups A1 and 2, A4 and 5, A11 and 15 and A17 and 18. This approach finally resulted in 19 differently sized subgroups of family A (Table 2) that were further subjected to the more reliable maximum-likelihood and quartet-puzzling algorithms. Maximum-likelihood approaches calculate the probability of the observed data assuming that it has evolved in accordance with a chosen evolutionary model. Phylogenies are then inferred by finding trees and parameters that

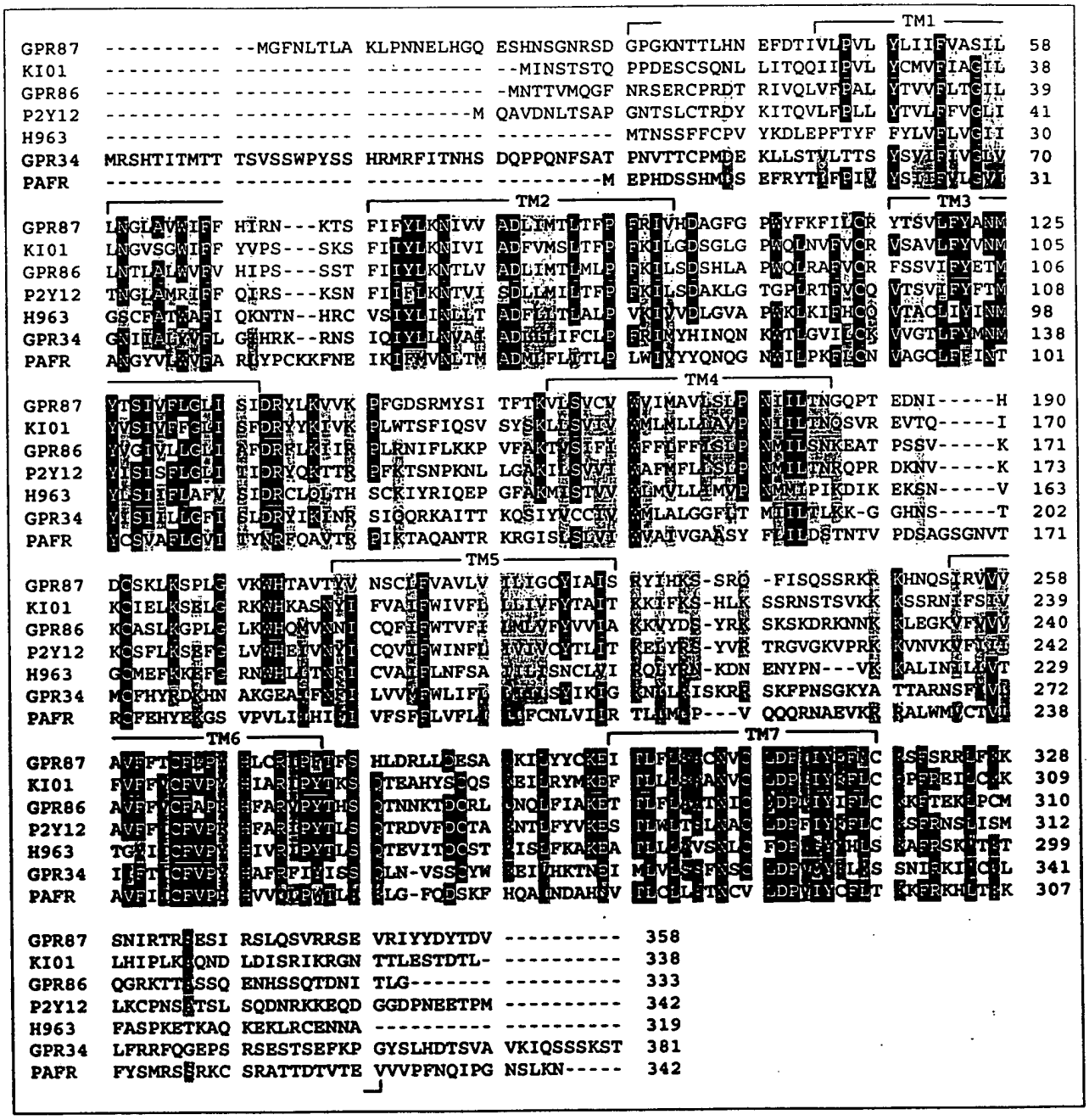


Figure 1

An example multiple sequence alignment of seven receptors. Protein sequences of GPR87, KI01, GPR86, P2Y12, H963, GPR34 and PAFR belonging to subgroup 12 were aligned with ClustalX and modified by deleting the extremely variable amino termini upstream of the first transmembrane domain and carboxyl termini downstream of the seventh transmembrane domain as indicated. Identical amino-acid residues in all aligned sequences are shaded in black and similar residues in gray. Transmembrane (TM) domains identified by the TMpred program are indicated.

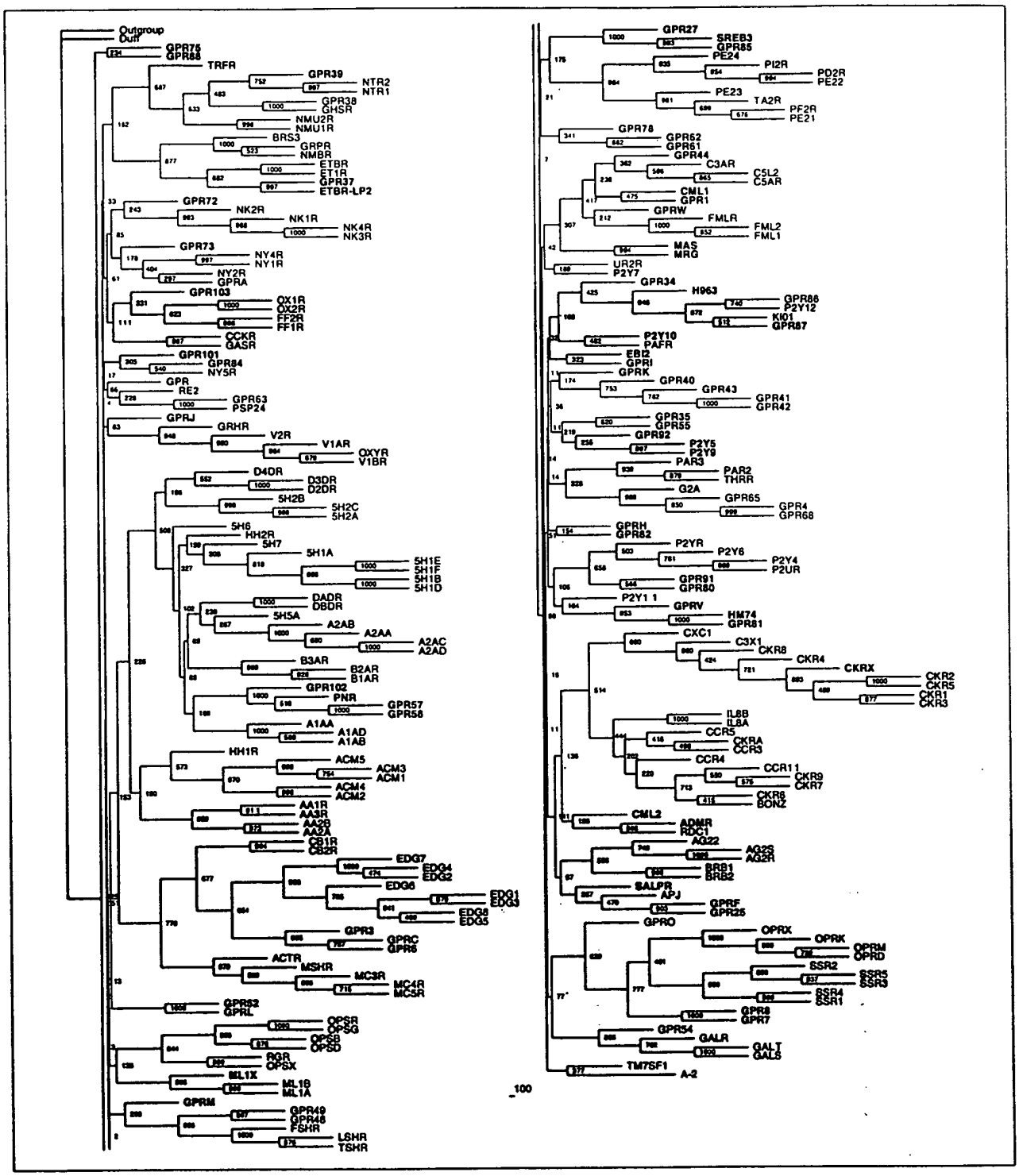


Figure 2

Neighbor-joining tree of the rhodopsin receptor-like family A inferred from the multiple sequence alignment using PHYLIP 3.6. Support values for each internal branch were obtained by 1,000 bootstrap steps, and are indicated. Pairwise distances were determined with PROTDIST and the JTT substitution frequency matrix. The tree was calculated with NEIGHBOR using standard parameters and rooted with the distant, though related, family-B receptor GPRC5B as the outgroup. The consensus tree of all bootstrapped sequences was obtained with CONSENSE. Orphan receptors are shown in bold. Scale bar indicates the branch length of 100 substitutions per site.

Table 2

Receptor subgroups derived from a combination of neighbor-joining and BLASTP results

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
C3X1	**ADMR**	AG22	**GPR7**	GALR	FF1R	BRS3	C3AR	**GPR72**	FSHR	**GPR40**
CKR1	BONZO	AG2R	**GPR8**	GALS	FF2R	ET1R	C5AR	**GPR73**	GPR48	**GPR41**
CKR2	CCR11	AG2S	OPRD	GALT	GASR	ETBR	C5L2	**(GPR75)**	GPR49	**GPR42**
CKR3	CCR3	APJ	OPRK	GPR54	**(GPR)**	ETBR-LP2	**CML1**	GPRA	LSHR	**GPR43**
CKR4	CCR4	BRB1	OPRM	GPRO	**GPR103**	GHSR	FML1	**GPRJ**	TSHR	**GPR80**
CKR5	CCR5	BRB2	OPRX	P2Y7	**(GPRM)**	**GPR37**	FML2	ML1A		**GPR81**
CKR8	CKR6	**GPR25**	SSR1	**SALPR**	GRHR	GPR38	FMLR	ML1B		**GPR82**
CKRX	CKR7	**GPRF**	SSR2	UR2R	OX1R	**GPR39**	**GPR1**	**MLIX**		**GPR91**
CXC1	CKR9		SSR3		OX2R	GRPR	**GPR44**	NK1R		**GPRV**
(TM7SF1)	CKRA		SSR4		OXYR	NMBR	**GPRW**	NK2R		**HM74**
	CML2		SSR5		V1AR	NMU1R	**(MAS)**	NK3R		P2UR
	(DUFF)				V1BR	NMU2R	**(MRG)**	NK4R		P2Y11
	IL8A				V2R	NTR1		NY1R		P2Y4
	IL8B					NTR2		NY2R		P2Y6
	RDC1					TRFR		NY4R		P2YR
								NY5R		

A12	A13	A14	A15	A16	A17	A18	A19	B	C
GPR34	ACTR	PD2R	**EB12**	OPSB	5H2A	AA1R	5H1A	BAI1	CASR
GPR86	CB1R	PE21	G2A	OPSD	5H2B	AA2A	5H1B	BAI2	GBR1
GPR87	CB2R	PE22	**GPR35**	OPSG	5H2C	AA2B	5H1D	BAI3	GBR2
H963	EDG1	PE23	GPR4	OPSR	5H6	AA3R	5H1E	CALR	**GPRC5B**
KI01	EDG2	PE24	**GPR55**	OPSX	A1AA	ACM1	5H1F	CD97	**GPRC5C**
P2Y12	EDG3	PF2R	GPR65	RGR	A1AB	ACM2	5H5A	CGRR	**GPRC5D**
PAFR	EDG4	PI2R	GPR68		A1AD	ACM3	5H7	CRF1	MGR1
	EDG5	TA2R	**GPR92**		A2AA	ACM4		CRF2	MGR2
	EDG6		**GPRH**		A2AB	ACM5		**EMR1**	MGR3
	EDG7		**GPRI**		A2AC	**GPR101**		**EMR2**	MGR4
	EDG8		**GPRK**		A2AD	**GPR27**		**EMR3**	MGR5
	GPR3		**P2Y10**		B1AR	**GPR52**		GIPR	MGR6
	GPR6		**P2Y5**		B2AR	**GPR61**		GLPR	MGR7
	GPRC		**P2Y9**		B3AR	**GPR62**		GLR	MGR8
	MC3R		PAR2		D2DR	**GPR63**		GPL2	
	MC4R		PAR3		D3DR	**GPR78**		**GPR56**	
	MC5R		THRR		D4DR	**GPR84**		GRFR	
	MSHR				DADR	**GPR85**		PACR	
					DBDR	**(GPR88)**		PTR2	
					GPR102	**GPRL**		PTRR	
					GPR57	HH1R		SCRC	
					GPR58	**PSP24**		VIPR	
					HH2R	**RE2**		VIPS	
					PNR	**SREB3**			

Very distantly related receptors that are possibly not phylogenetically related are shown in brackets. Orphan receptors are shown in bold.

yield the highest likelihood. Maximum-likelihood approaches tend to outperform alternative methods such as parsimony or distance-based methods. The main advantage is the application of a well defined model of sequence evolution to a given dataset [26]. Maximum likelihood is the estimation method least affected by sampling error and tends to be robust to many violations of the assumptions in the evolutionary model. The methods are statistically well founded, evaluate different tree topologies and use all sequence information available [27,28]. Because of their smaller size, families B and C could be subjected to these methods without prior subgrouping. This resulted in 19 phylogenetic trees, comprising 241 receptors for family A (Figures 3-6), one tree from 23 sequences for family B and one tree from 14 sequences for family C (Figure 7). Family-A trees were rooted with the human family-B receptor GPRC5B and families B and C with family-A receptor 5H1A. The sequence used to root the tree (the outgroup) is supposed to be a distant, though related, sequence. In some of our groups, the phylogenetic trees could not be fully resolved. This could be due to either very similar or very distant sequences. In both cases the phylogenetic signal is too weak to resolve the tree [29]. Several receptors (for example, TM7SF1, DUFF, GPR, GPRM, GPR75, GPR88, MAS and MRG) were found to be only distantly related to other known receptors used in our analysis. A possible explanation could be the previously proposed convergent evolution of this large protein family, meaning that these receptors have acquired the compelling similarity in their overall structures as a result of functional need, not phylogenetic relationship. The lack of significant sequence similarity among the different GPCR families favors this assumption [30-32]. Other explanations for the lack of significant sequence similarities might be an extraordinary divergence (genetic drift) or technical problems of the sequence-analysis methods used in analyzing polytopic membrane proteins or large protein families [33].

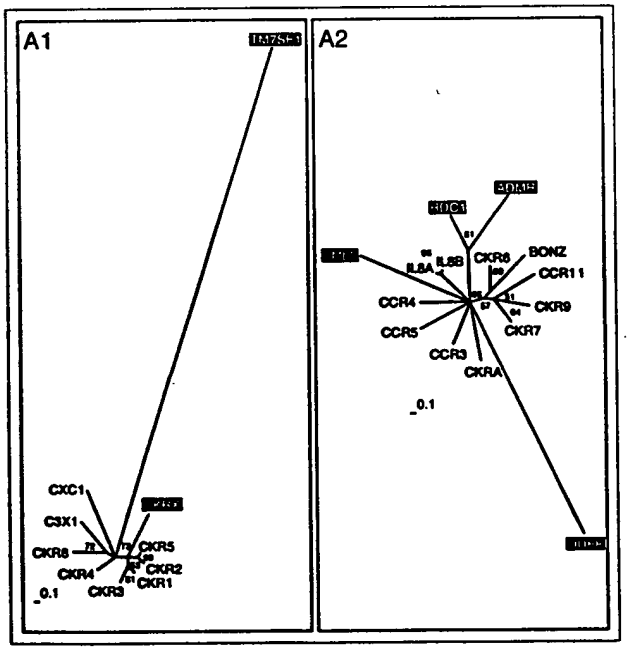


Figure 3
Chemokine receptors (subgroups A1 and A2). Phylogenetic trees of the subgroups were inferred using Puzzle 5.0 corrected by the JTT substitution frequency matrix. Quartet-puzzling support percentage values from 10,000 puzzling steps are shown. The scale bars indicate a maximum likelihood branch length of 0.1 inferred substitutions per site. Orphan receptors are shaded.

Receptor family A subgroups

In contrast to the subfamilies presented in GPCRDB [34], a database widely used in the field, our grouping shows the orphan receptors within their respective subgroup and their relationship to receptors with known ligands. In addition, our method sometimes resulted in subgroups with members whose ligands belong to different substance classes. These results are discussed in more detail below.

Chemokine receptors

Groups A1 and A2 comprise the chemokine receptors (Figure 3). The chemokine ligand superfamily is defined by four conserved cysteines that form two disulfide bonds, and can be structurally subdivided into two major branches based on the spacing of the first cysteine pair. Chemokines in which these residues are adjacent form the CC subfamily (corresponding to the SWISS-PROT CKR nomenclature used here), and those separated by a single amino acid comprise the CXC subfamily (here CCR and IL8R; for a review see [35]). We had to divide the whole subfamily into two groups to perform a detailed phylogenetic analysis. This subgrouping produced the same dichotomy, as suggested by the two-ligand motifs, as another example of the parallel evolution of receptors and ligands. Similar results describing this parallel evolution were found previously using a different computational approach [36].

Group A1 mainly comprises the CC family. We hypothesize that the orphan receptor CKRX, which constitutes a separate branch related to CKR1, 2, 3 and 5, might also bind a CC ligand. In contrast, TM7SF1 in this group seems to be only distantly, if at all, related to family-A receptors. It was grouped according to BLASTP results, where a misleading local alignment of approximately 20 amino acids placed it in the vicinity of the chemokine receptors. Group A2 is more heterogeneous and comprises receptors for CC and CXC ligands, as well as an orphan receptor (ADMR) previously thought to bind the peptide adrenomedullin. Adrenomedullin has now been shown to bind a family-B receptor and is discussed further below. The orphan receptor RDC1 in group A2 was first believed to be a receptor for vasointestinal peptide VIP [37], a notion not supported by phylogeny and later dismissed by experimental data [38]. Our results place it closer to the ADMR receptor than to the

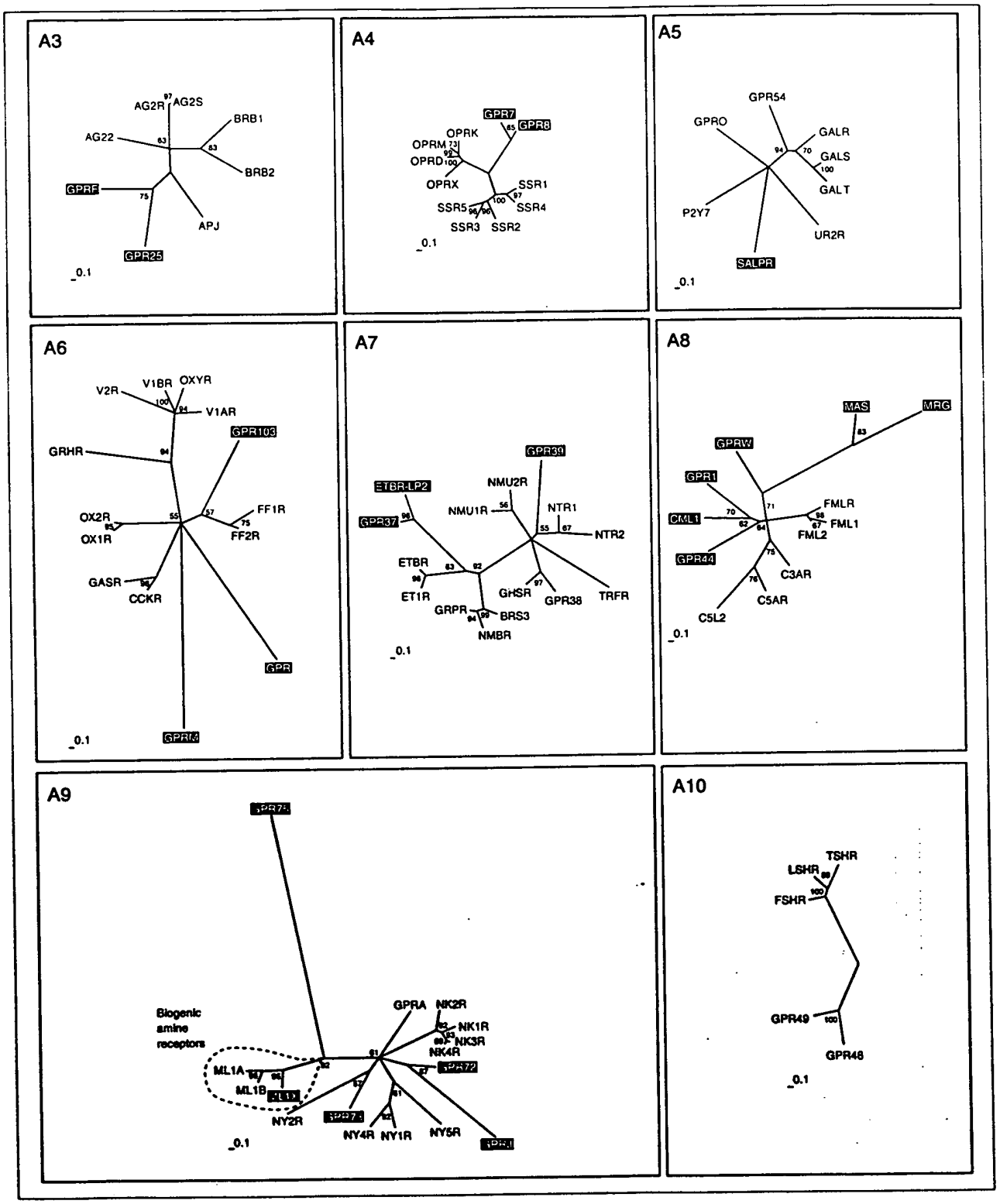


Figure 4
Peptide receptors (subgroups A3-A10). The scale bar indicates a maximum likelihood branch length of 0.1 inferred substitutions per site. Orphan receptors are shaded. For method see Figure 2.

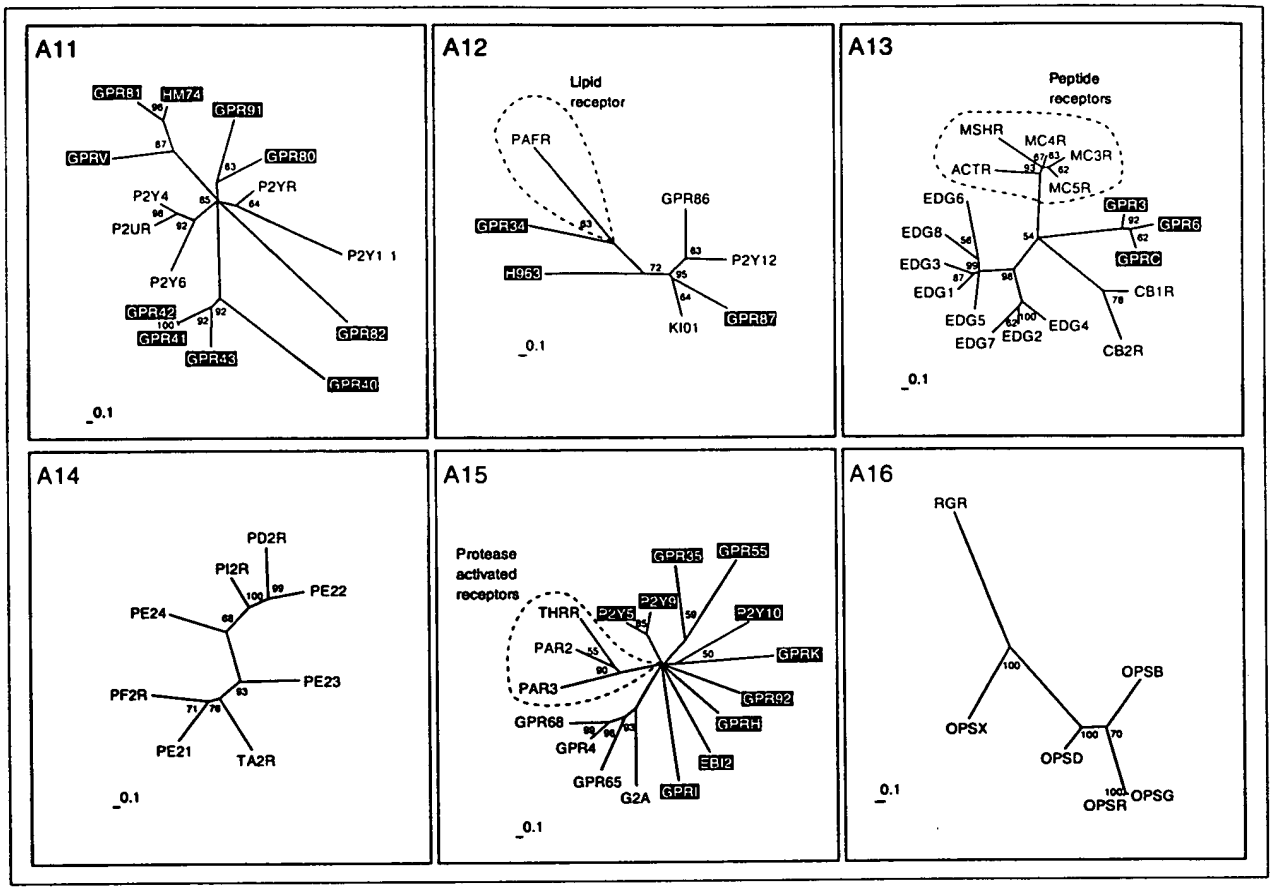


Figure 5
Nucleotide and lipid receptors (subgroups A11-A16). The scale bar indicates a maximum-likelihood branch length of 0.1 inferred substitutions per site. Orphan receptors are shaded. For method see Figure 2.

typical chemokine receptors. CML2 is a typical, but distant, member of the chemokine receptor family. The DUFF receptor (the Duffy antigen) is also very distantly related and was only grouped into A2 by BLASTP results.

Peptide receptors

Group A3 consists of receptors for the small peptides angiotensin (8 amino acids), bradykinin (9 amino acids) and apelin (Figure 4). Four forms of apelin (12, 13, 17 and 36 amino acids) have been described, but only those of 12 and 13 amino acids bind in nanomolar concentrations [39]. The orphan receptors GPRF and GPR25 in this group are related as closely to the apelin receptor APJ as to the angiotensin or bradykinin receptors, and might also bind small peptides. GPRF acts as a co-receptor for the human immunodeficiency virus (HIV) [40], like the APJ receptor [41], which further hints at structural homology of the two ligands. Opioid and somatostatin receptors make up group A4. Both somatostatin and opioid peptides are derived from the processing of larger precursors. The somatostatins are cyclic peptides of 14 and 28 amino acids. The opioid precursors preproenkephalin, preprodynorphin, prepro-opiomelanocortin and prepronociceptin display a strikingly similar general organization and a conserved amino-terminal region that contains six cysteines, probably involved in disulfide bond formation.

The processed neuropeptides, in contrast, are less similar to each other. It could be speculated that the receptors first bound the precursors themselves, and that the diversity derived from processing is evolutionarily new. Processing prepronociceptin gives rise to two evolutionarily conserved peptides besides orphanin FQ, the ligand for OPRX. It has not been reported whether these peptides bind to the orphan receptors GPR7 and GPR8, which constitute a new branch related to the opioid receptors.

In group A5 we find three receptors that bind the 30-amino-acid peptide galanin, and related to these the GPR54 receptor, which is activated by the 54-, 14-, and 13-amino-acid peptides derived from the product of KiSS-1, a metastasis

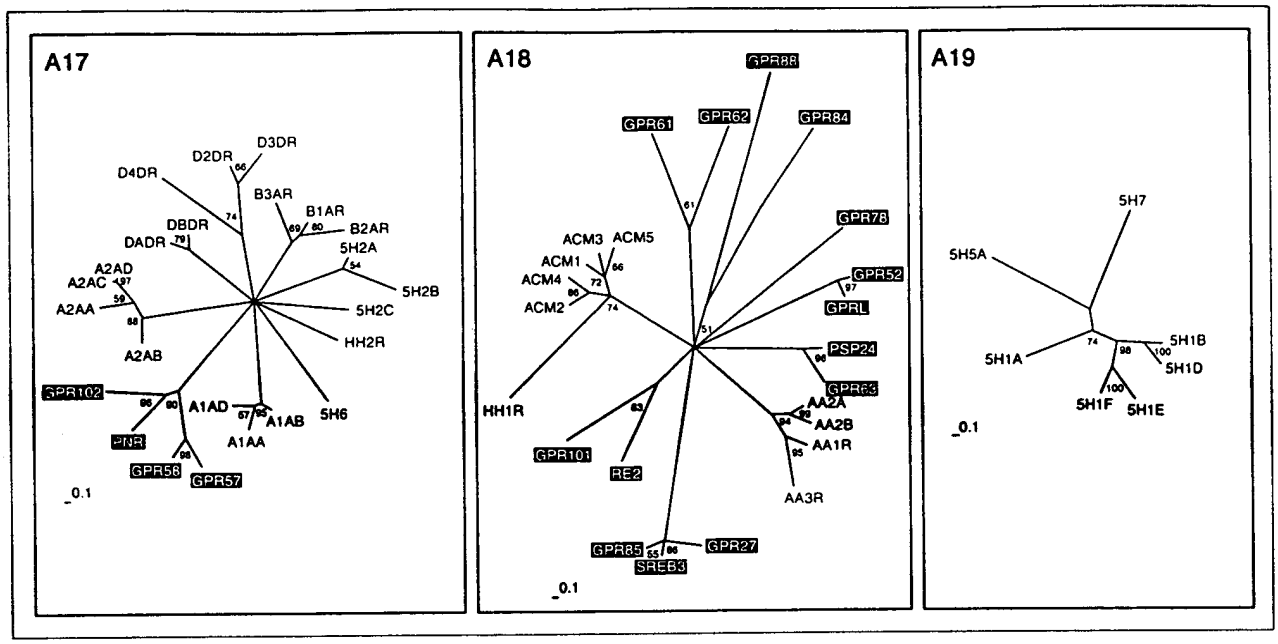


Figure 6
Biogenic amine receptors (subgroups A17-A19). The scale bar indicates a maximum-likelihood branch length of 0.1 inferred substitutions per site. Orphan receptors are shaded. For method see Figure 2.

suppressor gene for melanoma cells. These kisspeptins all share a common RF-amide caboxyl terminus. Although only distantly related to each other, both GPRO (melanin-concentrating hormone) and UR2R (urotensin II peptide) bind cyclic peptides originally isolated from fish. Similarly distant is the orphan receptor SALPR, which shares sequence similarity with somatostatin (A4) and angiotensin (A3) receptors, but subgrouping of groups A4 and 5 by neighbor joining led to its placement in group 5. SALPR does not bind somatostatin or angiotensin ligands [42], but could bind another cyclic peptide. The P2Y7 receptor in group A5 does not bind nucleotides [43], as suggested by the name, but was published as a receptor for the lipid leukotriene B4 [44], a notion not supported by phylogeny. In addition, two new leukotriene receptors - CLT1 and CLT2 - have been cloned and characterized during the preparation of this manuscript [45,46] and were found to be unrelated to P2Y7.

Group A6 is again composed solely of receptors for peptide ligands. The orphan receptor GPR103 is related to the neuropeptide FF receptors that bind two amidated mammalian neuropeptides - NPAF (A-18-F-amide) and NPFF (F-8-F-amide), also known as morphine-modulating peptides. These peptides, which may also be the ligand for GPR103, are members of a large family of neuropeptides related to the molluscan cardioexcitatory neuropeptide (FMRF-amide, Phe-Met-Arg-Phe-amide). The orphan receptors GPRM and GPR in group A6 are most probably also peptide receptors, but are only very distantly related to the others and show no relationship to receptors with known ligands. Group A7 is also composed of receptors for peptide ligands: neuromedin, neurotensin, motilin, endothelin, bombesin and the releasing hormones for growth hormone and thyrotropin. GPR39 might bind a small peptide ligand like the closely related neurotensin receptors NTR1 and 2, which binds a 13-amino-acid peptide derived from a larger precursor protein. GPR37 and ETBR-LP2 are related to each other and branch off the endothelin receptors that bind characteristic bicyclic peptides of 21 amino acids containing four cysteines linked by two disulfide bonds.

Group A8 has two branches with receptors with known ligands. These receptors bind the structurally diverse but functionally related chemotactic substances *N*-formylmethionyl and the anaphylatoxic complement factors. The *N*-formylmethionyl ligands are small hydrophilic peptides of bacterial origin, but recently a number of new peptide agonists have been identified that selectively activate the high-affinity fMLF receptor FPR and/or its low-affinity variant FPRL1. These agonists include peptide domains derived from the envelope proteins of HIV type 1 and at least three amyloidogenic polypeptides, the human acute-phase protein serum amyloid A, the 42-amino-acid form of beta-amyloid peptide and a 21-amino-acid fragment of the human prion protein. Furthermore, a cleavage fragment of neutrophil granule-derived bactericidal cathelicidin, LL-37, is also a chemotactic agonist for FPRL1 (for a review see [47]). The complement factors C3a and C5a are large but highly

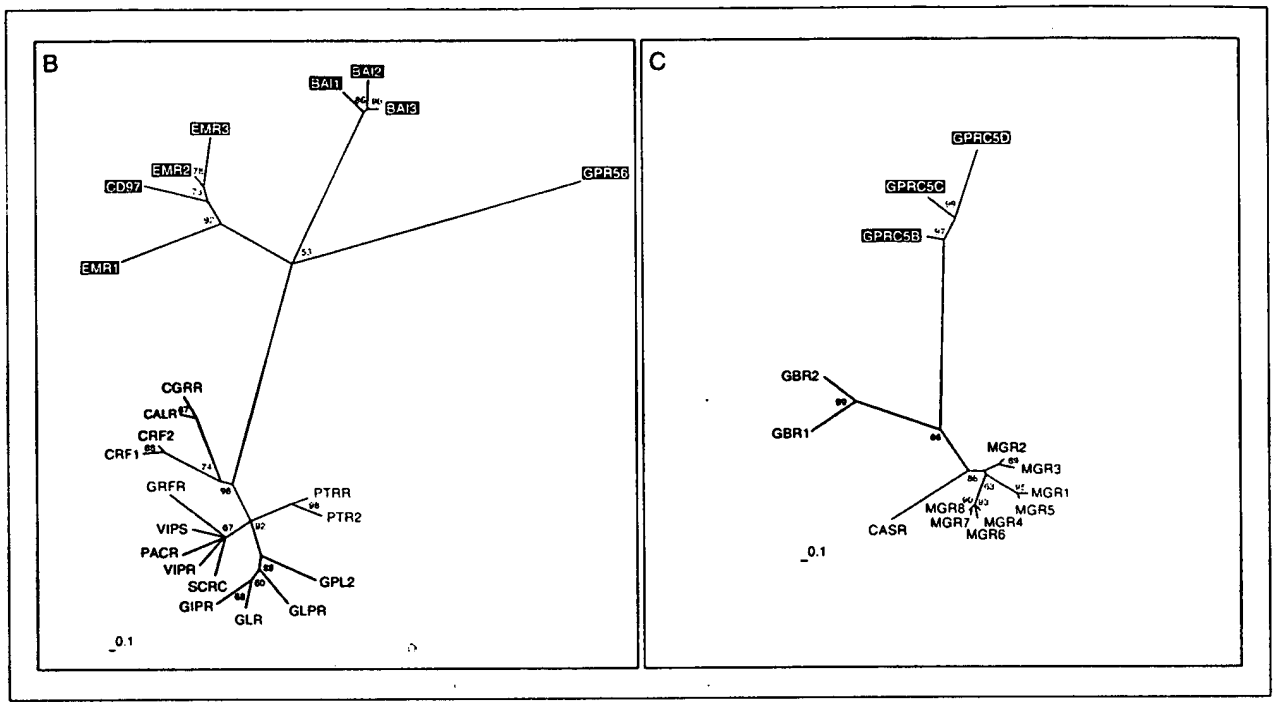


Figure 7
Families B and C of the G-protein-coupled receptors (GPRCs). Phylogenetic trees of families B and C were inferred using Puzzle 5.0 corrected by the JTT substitution frequency matrix. Quartet-puzzling support percentage values from 10,000 puzzling steps are shown. The scale bar indicates a maximum likelihood branch length of 0.1 inferred substitutions per site. Orphan receptors are shaded.

hydrophilic proteins with a mainly alpha-helical structure held together by three disulfide bridges. C5a is rapidly desarginated to the less potent derivative C5adR74, which is the ligand for the C5L2 receptor. The orphan receptors GPR1, CML1 and GPR44 all cluster, and constitute a separate branch as distant as the other two branches. No prediction of the possible structure of the ligands for these receptors can be derived from this tree, but maybe they will function as chemotactic peptides. This could at least hint at leukocytes or inflamed tissue as a possible source for these ligands. The receptor GPRW constitutes its own branch, not as distant to the main group as the MAS oncogene product and the related receptor MRG, which are only very distantly related to the group.

All receptors in group A9 with known ligands bind peptides, except for a side branch consisting of receptors for the biogenic amine melatonin. The orphan receptor ML1X is closely related to melatonin receptors ML1A and B, but apparently does not bind melatonin [48]. GPR73 is related to the neuropeptide Y (NPY) receptor NY2R which mainly binds the pancreatic peptide YY of 36 amino acids, and these two are placed together on a branch distinct from the NPY receptors NY4R and NY1R. GPR73 does not bind the NPY ligand family [49], but possibly a similar large peptide ligand. The orphan receptors GPR72 and GPRJ constitute a new subgroup that most probably bind related peptide ligands. GPR72 does not bind a NPY ligand [49]. GPR75 is only very distantly related to the whole A9 group. The receptors for the glycoprotein hormones thyroid-stimulating hormone (TSH), luteinizing hormone (LSH) and follicle-stimulating hormone (FSH) make up Group A10. GPR48 and 49 are very similar in their overall structure, with long amino termini, but their relationship is also evident in the neighbor-joining tree constructed from alignments without amino and carboxyl termini. It has been recently shown that these receptors mediate the action of relaxin, a peptide hormone of the insulin-like growth factor family secreted by the corpus luteum during pregnancy [50].

Nucleotide and lipid receptors

The receptors with known ligands in group A11 are the P2Y receptors, which bind pyrimidine as well as purine nucleotides (Figure 5). Several orphan receptors constitute new clusters. GPR80 and GPR91 are distantly related to each other and relatively close to the P2Y receptors. GPR80 is the closest relative of the newly identified CLT2 receptor for leukotrienes as judged by BLASTP results. GPR81, HM74 and GPRV and GPR 40-43 belong to branches only distantly related to P2Y receptors. Within these potential new subfamilies, GPR41-43, GPR81 and HM74 are more closely related to each other than to GPR40 (for GPR41-43) and GPRV (for GPR81 and HM74).

In group A12, the platelet-activated receptor, a lipid receptor and receptors activated by nucleotides mingle, but are found on different side branches. The orphan receptor GPR87 is closely related to the receptor for UDP-glucose KI01 and to the ADP-binding receptors P2Y12 and GPR86. We assume that this receptor might also bind UDP-glucose or another modified nucleotide. GPR34 is distantly related to the platelet-activating factor (PAF) receptor; it was not activated by available lipid ligands [51], but might nevertheless bind a lipid ligand. Group A13 contains both peptide and lipid receptors but they make up different branches. The peptide branch binds peptides derived from the processing of pro-opiomelanocortin that gives rise to peptides of between 12 and 36 amino acids. The EDG and cannabinoid receptors constitute clusters, and one cluster distinct from the other three consists of the orphan receptors GPR3, GPR6 and GPRC, which have been grouped closer to the lipid EDG receptors in the overall neighbor-joining tree (Figure 2). This information helped to identify a phospholipid ligand for GPRC (H. Chica Schaller, personal communication).

The receptors in group A14 all bind ligands derived from arachidonic acid by the action of cyclooxygenase. These receptors for lipid-derived autacoids or prostanoids comprise receptors for the prostaglandins and thromboxanes. There are no orphan receptors in this group. Group A15 is a very heterogenous group composed of receptors for the lipids sphingosylphosphorylcholine (SPC), lysophosphatidylcholine (LPC) and psychosine, and receptors activated by proteases. GPR4 and GPR68 both bind SPC, like the EDG receptor branch consisting of the EDG1, 3, 6 and 8 receptors in A13, but are not closely related. Protease-activated receptors become activated by a part of the former amino terminus cleaved by the protease. The new amino terminus then functions as a tethered ligand and activates the receptor. This can be mimicked by very small peptides derived from this ligand; such receptors should therefore rather resemble peptide receptors. The orphans P2Y5, P2Y9 and P2Y10 receptors were not placed in group 11 and 12 like most P2Y receptors, but in group A15, supporting the fact that they were misnamed. P2Y5 and P2Y9 do not bind nucleotides [52,53], but this has not been shown yet for P2Y10. All other orphan receptors in this group, with the exception of GPR35 and GPR55 which cluster together, are as distantly related to each other as to the receptors with known ligands. Group A16 contains the opsins, receptors that are activated by isoprenoid ligands, and no orphan receptors.

Biogenic amine receptors

Some serotonin receptors and receptors for the biogenic amines adrenaline, dopamine and histamine are all placed on different branches in group A17 (Figure 6). An additional branch consists of the orphan receptors GPR102, PNR, GPR57 and GPR58, which are as distantly related to the others as, for example, is the alpha-adrenergic receptor branch. PNR and GPR58 expressed in COS cells did not bind various serotonin receptor-specific ligands [54]. Their ligands might be small molecules with similar properties. Group A18 is very heterogeneous and consists of receptors for the biogenic amines acetylcholine and adenosine, and the HH1R receptor for histamine, as well as many distantly related orphan GPCRs. GPR63 is closely related to the orphan receptor PSP24. The *Xenopus laevis* homolog of this receptor binds LPA [55]. GPR101 and RE2, GPRL and GPR52, and GPR61 and GPR62 constitute their own subgroups. In particular, the SREB1-3 cluster (GPR85, GPR27 and SREB3) makes up its own family, with only a distant relationship to other GPCRs in this group. No orphan receptors are found in group A19, which consists entirely of serotonin receptors distinct from those in A17.

During the preparation of this manuscript several new family-A receptors that could not be fitted into our analysis were identified. These comprise 15 new receptors distinct from the classical biogenic amine receptors that apparently bind the trace amines tyramine, β-phenylethylamine, tryptamine and octopamine [56]. In addition, a new subfamily of GPCRs related to the *mas* oncogene and uniquely expressed in small nociceptive sensory neurons were shown to be the receptors for a number of enkephalin fragments [57].

Receptor families B and C

Family B (Figure 7) was named after the secretin receptor. Yet proteins showing homology to this receptor make up only one of four distantly related subgroups. The receptors EMR1, EMR2 and EMR3, and the CD97 surface antigen, all have several epidermal growth factor (EGF)-like domains in the extracellular amino terminus. They constitute their own cluster only distantly related to the rest of the family. The same applies to the brain-specific angiogenesis inhibitor family BAI1-3. GPR56 was assigned to family B because it shows the typical signature [58], but is so far the only one of its kind. So far no non-protein ligand has been identified as a ligand for family-B receptors. Astonishingly, one family-B receptor, namely the CGRP receptor, requires coexpression with single transmembrane receptor activity-modifying proteins (RAMP1-3) for ligand binding and signal transduction [59]. Coexpression of different RAMPs results in binding of different cyclic peptide ligands such as adrenomedullin, amylin or the calcitonin gene-related peptide (for a review see [60]). This could further complicate the identification of the cognate ligands for these family-B orphan receptors, but we assume that they will also bind large peptide ligands. In family C (Figure 7), the metabotropic glutamate receptors MGR1-8 bind the small molecule glutamate, the CASR receptor senses extracellular calcium concentration, and receptors GBR1-2 bind the small molecule gamma-amino butyric acid (GABA). GPRC5B, C and D constitute their own subgroup with no closer relationship to the other members, but might also bind small molecules.

Conclusions

In this work, we calculated the phylogenetic distances of 277 human GPCRs and show the relationship of orphan receptors to receptors for known ligands with support values for each branch. We then grouped orphan receptors and receptors with known ligands into 19 subgroups that sometimes differ from previous classifications. Three subgroups are composed of receptors for ligands that belong to different substance classes; for example, in group A12, lipid receptors and receptors activated by nucleotides mingle, and in groups A13 and A15, peptide and lipid receptors. In both subgroups the receptors binding ligands of different substance classes make up different branches. We hope that this approach proves valuable for identifying the natural ligands of orphan receptors, as related receptors have previously been shown to have ligands with similar structural features.

Materials and methods

Sequence database mining

A database search excluding olfactory and gustatory receptors identified the amino-acid sequences of 281 human GPCRs. Only sequences annotated as GPCRs in the following databases were used: NCBI [61], SWISS-PROT [62], EMBL [63] and GPCRDB [34,64]. Receptors without published ligands in PubMed [65] were defined as orphan GPCRs.

Multiple sequence alignments

Multiple protein sequences were aligned with ClustalX 1.81 [66]. Pairwise alignment parameters were set as: slow/accurate alignment; gap opening penalty 10; gap extension penalty 0.10; protein weight matrix BLOSUM 30. Multiple alignment parameters were set as: gap opening penalty 10; gap extension penalty 0.05; delay divergent sequences 35%; protein weight matrix BLOSUM series [67]. The alignments were modified by deleting the extremely variable amino termini upstream of the first transmembrane domain and carboxyl termini downstream of the seventh transmembrane domain. Alignment editing and shading was done using BioEdit Sequence Alignment Editor [68] and GeneDoc Multiple Sequence Alignment Editor [69]. Transmembrane domains were identified using the TMpred program [70] and, where available, data from the original publication [71].

Clustering of subgroups

An overall phylogenetic tree of family A was inferred from the multiple sequence alignment with PHYLIP 3.6 [72]. Bootstrapping was performed 1,000 times using SEQBOOT to obtain support values for each internal branch. Pairwise distances were determined with PROTDIST and the JTT substitution frequency matrix [73]. Neighbor-joining phylogenetic trees [21] were calculated with NEIGHBOR using standard parameters. The human GPRC5B receptor belonging to family B was used as outgroup for family A. The outgroup sequence is supposed to be a distant, though related, sequence and is used to root the tree. The majority-rule consensus trees of all bootstrapped sequences were obtained with the program CONSENSE. Representations of the calculated trees were constructed with TreeView [74]. Clusters with bootstrap values greater than 50% were defined as confirmed subgroups, and sequences with lower values added to these subgroups according to their sequence similarity in the alignment as judged by visual inspection and the results of pairwise local alignments with all other sequences by BLASTP [25]. The *p*-value was used as a measure of similarity.

Quartet-puzzling trees

Multiple protein sequence alignments of these new subgroups were created as described above. Phylogenetic trees were inferred from these alignments using Puzzle 5.0 [75] to calculate maximum-likelihood distances corrected by the JTT substitution-frequency matrix [73] with amino-acid usage estimated from the data, site-to-site rate variation modeled on a gamma distribution with eight rate categories plus invariant sites, and the shape parameter estimated from the data. The human GPRC5B receptor of family B was used as an outgroup for family A. The human 5H1A receptor of family A was used as an outgroup for families B and C (the outgroups are not shown in the figures here). Quartet-puzzling (QP) trees were constructed with the described settings and 10,000 puzzling steps to obtain support values (QP reliability) for each internal branch. The program Puzzle 5.0 was used in a parallelized version (ppuzzle) with a message-passing interface (MPI) implementation on a HP 9000 N-Class Enterprise Server Cluster consisting of five HP 9000 N-Class shared-memory multiprocessor systems with eight PA-RISC 8600 (552 MHz) processors each. Representations of the quartet-puzzling trees were constructed with TreeView [74].

Additional data files

Additional data files available with the online verson of this paper include a data table with names, synonyms and accession numbers of all GPCRs, and the BLASTP results of all GPCRs (full-length sequences and sequences without amino or carboxyl termini).

Acknowledgements

The DFG Graduiertenkolleg 255, the Dr Kurt und Irmgard Meister-Stiftung and the Hamburgische Wissenschaftliche Gesellschaft, supported this study. We appreciate the help of Chica Schaller in finding additional sequences and of Andreas Schuldei in reconfiguring ppuzzle and using MPI. Klaus Martens and his colleagues at the computing center of the Technical University Hamburg-Harburg provided an account at the HP N-Class Enterprise Server Cluster and helped us to use the software environment.

References

1. Gether U: **Uncovering molecular mechanisms involved in activation of G protein-coupled receptors.** *Endocr Rev* 2000, 21:90-113.
2. Libert F, Parmentier M, Lefort A, Dinsart C, Van Sande J, Maenhaut C, Simons MJ, Dumont JE, Vassart G: **Selective amplification and cloning of four new members of the G protein-coupled receptor family.** *Science* 1989, 244:569-572.

3. Methner A, Hermey G, Schinke B, Hermans-Borgmeyer I: **A novel G protein-coupled receptor with homology to neuropeptide and chemoattractant receptors expressed during bone development.** *Biochem Biophys Res Commun* 1997, **233**:336-342.
4. Lee D, George S, Evans J, Lynch K, O'Dowd B: **Orphan G protein-coupled receptors in the CNS.** *Curr Opin Pharmacol* 2001, 1:31-39.
5. Reinscheid RK, Nothacker HP, Bourson A, Ardati A, Henningsen RA, Bunzow JR, Grandy DK, Langen H, Monsma FJ Jr, Civelli O: **Orphanin FQ: a neuropeptide that activates an opioidlike G protein-coupled receptor.** *Science* 1995, **270**:792-794.
6. Civelli O, Nothacker HP, Saito Y, Wang Z, Lin SH, Reinscheid RK: **Novel neurotransmitters as natural ligands of orphan G-protein-coupled receptors.** *Trends Neurosci* 2001, 24:230-237.
7. Drews II: **Drug discovery today - and tomorrow.** *Drug Discov Today* 2000, 5:2-4.
8. Lee MJ, Van Brocklyn JR, Thangada S, Liu CH, Hand AR, Menzeleev R, Spiegel S, Hla T: **Sphingosine-1-phosphate as a ligand for the G protein-coupled receptor EDG-1.** *Science* 1998, **279**:1552-1555.
9. Van Brocklyn JR, Tu Z, Edsall LC, Schmidt RR, Spiegel S: **Sphingosine 1-phosphate-induced cell rounding and neurite retraction are mediated by the G protein-coupled receptor H218.** *J Biol Chem* 1999, **274**:4626-4632.
10. Van Brocklyn JR, Graler MH, Bernhardt G, Hobson JP, Lipp M, Spiegel S: **Sphingosine-1-phosphate is a ligand for the G protein-coupled receptor EDG-6.** *Blood* 2000, **95**:2624-2629.
11. Im DS, Heise CE, Ancellin N, O'Dowd BF, Shei GJ, Heavens RP, Rigby MR, Hla T, Mandala S, McAllister G, *et al.*: **Characterization of a novel sphingosine 1-phosphate receptor, Edg-8.** *J Biol Chem* 2000, **275**:14281-14286.
12. Hecht JH, Weiner JA, Post SR, Chun J: **Ventricular zone gene-1 (vzg-1) encodes a lysophosphatidic acid receptor expressed in neurogenic regions of the developing cerebral cortex.** *J Cell Biol* 1996, **135**:1071-1083.
13. An S, Bleu T, Hallmark OG, Goetzl EJ: **Characterization of a novel subtype of human G protein-coupled receptor for lysophosphatidic acid.** *J Biol Chem* 1998, **273**:7906-7910.
14. Im DS, Heise CE, Harding MA, George SR, O'Dowd BF, Theodorescu D, Lynch KR: **Molecular cloning and characterization of a lysophosphatidic acid receptor, Edg-7, expressed in prostate.** *Mol Pharmacol* 2000, **57**:753-759.
15. Szekeres PG, Muir AI, Spinage LD, Miller JE, Butler SI, Smith A, Rennie GI, Murdock PR, Fitzgerald LR, Wu H, *et al.*: **Neuromedin U is a potent agonist at the orphan G protein-coupled receptor FM3.** *J Biol Chem* 2000, **275**:20247-20250.
16. Zhu Y, Michalovich D, Wu H, Tan KB, Dytko GM, Mannan IJ, Boyce R, Alston J, Tierney LA, Li X, *et al.*: **Cloning, expression, and pharmacological characterization of a novel human histamine receptor.** *Mol Pharmacol* 2001, **59**:434-441.
17. Communi D, Gonzalez NS, Detheux M, Brezillon S, Lannoy V, Parmentier M, Boeynaems JM: **Identification of a novel human ADP receptor coupled to G(i).** *J Biol Chem* 2001, **276**:41479-41485.
18. Chambers JK, Macdonald LE, Sarau HM, Ames RS, Freeman K, Foley JJ, Zhu Y, McLaughlin MM, Murdock P, McMillan L, *et al.*: **A G protein-coupled receptor for UDP-glucose.** *J Biol Chem* 2000, **275**:10767-10771.
19. Attwood TK, Findlay JB: **Fingerprinting G-protein-coupled receptors.** *Protein Eng* 1994, **7**:195-203.
20. Kolakowski LF Jr: **GCRDb: a G-protein-coupled receptor database.** *Receptors Channels* 1994, **2**:1-7.
21. Saitou N, Nei M: **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4**:406-425.
22. Page RD, Holmes EC: *Molecular Evolution: A Phylogenetic Approach.* Oxford: Blackwell Science; 1999: 218-225.
23. Kumar S, Gadagkar SR: **Efficiency of the neighbor-joining method in reconstructing deep and shallow evolutionary relationships in large phylogenies.** *J Mol Evol* 2000, **51**:544-553.
24. Fitch WM, Margoliash E: **Construction of phylogenetic trees.** *Science* 1967, **155**:279-284.
25. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410.
26. Felsenstein J: **Phylogenies from molecular sequences: inference and reliability.** *Annu Rev Genet* 1988, **22**:521-565.
27. Strimmer K, Goldman N, von Haeseler A: *Bayesian Probabilities and Quartet Puzzling. Mol Biol Evol* 1997, **14**:210-211.
28. Strimmer K, von Haeseler A: **Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies.** *Mol Biol Evol* 1996, **13**:964-211.
29. Strimmer K, Robertson DL: **Inference and applications of molecular phylogenies: an introductory guide.** In *The Internet for Molecular Biologists (Practical Approach Series)*, edited by Sansom C, Horton RM. Oxford: Oxford University Press; 2001.
30. Pin JP, Joly C, Heinemann SF, Bockaert J: **Domains involved in the specificity of G protein activation in phospholipase C-coupled metabotropic glutamate receptors.** *EMBO J* 1994, 13:342-348.
31. Probst WC, Snyder LA, Schuster DI, Brosius J, Sealfon SC: **Sequence alignment of the G-protein coupled receptor superfamily.** *DNA Cell Biol* 1992, 11:1-20.
32. Burbach JP, Meijer OC: **The structure of neuropeptide receptors.** *Eur J Pharmacol* 1992, **227**:1-18.
33. Graul R, Sadee W: **Evolutionary relationships among G protein-coupled receptors using a clustered database approach.** *AAPS PharmSci* 2001, **3**:E12.
34. Horn F, Weare J, Beukers MW, Horsch S, Bairoch A, Chen W, Edvardsen O, Campagne F, Vriend G: **GPCRDB: an information system for G protein-coupled receptors.** *Nucleic Acids Res* 1998, 26:275-279.
35. Cascieri MA, Springer MS: **The chemokine/chemokine-receptor family: potential and progress for therapeutic intervention.** *Curr Opin Chem Biol* 2000, 4:420-427.
36. Goh C, Bogan A, Joachimiak M, Walther D, Cohen F: **Co-evolution of proteins with their interaction partners.** *J Mol Biol* 2000, **299**:283-293.
37. Sreedharan SP, Robichon A, Peterson KE, Goetzl EJ: **Cloning and expression of the human vasoactive intestinal peptide receptor.** *Proc Natl Acad Sci USA* 1991, **88**:4986-4990.
38. Nagata S, Ishihara T, Robberecht P, Libert F, Parmentier M, Christophe J, Vassart G: **RDC1 may not be VIP receptor.** *Trends Pharmacol Sci* 1992, **13**:102-103.
39. Tatemoto K, Hosoya M, Habata Y, Fujii R, Kakegawa T, Zou MX, Kawamata Y, Fukusumi S, Hinuma S, Kitada C, *et al.*: **Isolation and characterization of a novel endogenous peptide ligand for the human APJ receptor.** *Biochem Biophys Res Commun* 1998, **251**:471-476.
40. Farzan M, Choe H, Martin K, Marcon L, Hofmann W, Karlsson G, Sun Y, Barrett P, Marchand N, Sullivan N, *et al.*: **Two orphan seven-transmembrane segment receptors which are expressed in CD4-positive cells support simian immunodeficiency virus infection.** *J Exp Med* 1997, **186**:405-411.
41. Choe H, Farzan M, Konkel M, Martin K, Sun Y, Marcon L, Cayabyab M, Berman M, Dorf ME, Gerard N, *et al.*: **The orphan seven-transmembrane receptor apj supports the entry of primary T-cell-line-tropic and dualtropic human immunodeficiency virus type 1.** *J Virol* 1998, **72**:6113-6118.
42. Matsumoto M, Kamohara M, Sugimoto T, Hidaka K, Takasaki J, Saito T, Okada M, Yamaguchi T, Furuichi K: **The novel G-protein coupled receptor SALPR shares sequence similarity with somatostatin and angiotensin receptors.** *Gene* 2000, **248**:183-189.
43. Herold CL, Li Q, Schachter JB, Harden TK, Nicholas RA: **Lack of nucleotide-promoted second messenger signaling responses in 1321N1 cells expressing the proposed P2Y receptor, p2y7.** *Biochem Biophys Res Commun* 1997, **235**:717-721.
44. Yokomizo T, Izumi T, Chang K, Takuwa Y, Shimizu T: **A G-protein-coupled receptor for leukotriene B4 that mediates chemotaxis.** *Nature* 1997, **387**:620-624.
45. Lynch KR, O'Neill GP, Liu Q, Im DS, Sawyer N, Metters KM, Coulombe N, Abramovitz M, Figueroa DJ, Zeng Z, *et al.*: **Characterization of the human cysteinyl leukotriene CysLT1 receptor.** *Nature* 1999, **399**:789-793.
46. Takasaki J, Kamohara M, Matsumoto M, Saito T, Sugimoto T, Ohishi T, Ishii H, Ota T, Nishikawa T, Kawai Y, *et al.*: **The molecular characterization and tissue distribution of the human cysteinyl leukotriene CysLT(2) receptor.** *Biochem Biophys Res Commun* 2000, **274**:316-322.
47. Le Y, Yang Y, Cui Y, Yazawa H, Gong W, Qiu C, Wang JM: **Receptors for chemotactic formyl peptides as pharmacological targets.** *Int Immunopharmacol* 2002, **2**:1-13.
48. Reppert SM, Weaver DR, Ebisawa T, Mahle CD, Kolakowski LF Jr: **Cloning of a melatonin-related receptor from human pituitary.** *FEBS Lett* 1996, **386**:219-224.

49. Parker R, Liu M, Eyre HJ, Copeland NG, Gilbert DJ, Crawford J, Sutherland GR, Jenkins NA, Herzog H: **Y-receptor-like genes GPR72 and GPR73: molecular cloning, genomic organisation and assignment to human chromosome 11q21.1 and 2p14 and mouse chromosome 9 and 6.** *Biochim Biophys Acta* 2000, 1491:369-375.
50. Hsu S, Nakabayashi K, Nishi S, Kumagai J, Kudo M, Sherwood O, Hsueh A: **Activation of orphan receptors by the hormone relaxin.** *Science* 2002, 295:671-674.
51. Schoneberg T, Schulz A, Grosse R, Schade R, Henklein P, Schultz G, Gudermann T: **A novel subgroup of class I G-protein-coupled receptors.** *Biochim Biophys Acta* 1999, 1446:57-70.
52. Li Q, Schachter JB, Harden TK, Nicholas RA: **The 6H1 orphan receptor, claimed to be the p2y5 receptor, does not mediate nucleotide-promoted second messenger responses.** *Biochem Biophys Res Commun* 1997, 236:455-460.
53. Janssens R, Boeynaems JM, Godart M, Communi D: **Cloning of a human heptahelical receptor closely related to the P2Y5 receptor.** *Biochem Biophys Res Commun* 1997, 236:106-112.
54. Lee DK, Lynch KR, Nguyen T, Im DS, Cheng R, Saldivia VR, Liu Y, Liu IS, Heng HH, Seeman P, *et al.*: **Cloning and characterization of additional members of the G protein- coupled receptor family.** *Biochim Biophys Acta* 2000, 1490:311-323.
55. Guo Z, Liliom K, Fischer DJ, Bathurst IC, Tomei LD, Kiefer MC, Tigyi G: **Molecular cloning of a high-affinity receptor for the growth factor-like lipid mediator lysophosphatidic acid from *Xenopus* oocytes.** *Proc Natl Acad Sci USA* 1996, 93:14367-14372.
56. Borowsky B, Adham N, Jones K, Raddatz R, Artymyshyn R, Ogozalek K, Durkin M, Lakhlani PP, Bonini JA, Pathirana S, *et al.*: **Trace amines: identification of a family of mammalian G protein-coupled receptors.** *Proc Natl Acad Sci USA* 2001, 98:8966-8971.
57. Lembo P, Grazzini E, Groblewski T, O'Donnell D, Roy M, Zhang J, Hoffert C, Cao J, Schmidt R, Pelletier M, *et al.*: **Proenkephalin A gene products activate a new family of sensory neuron-specific GPCRs.** *Nat Neurosci* 2002, 5:201-209.
58. Liu M, Parker RM, Darby K, Eyre HJ, Copeland NG, Crawford J, Gilbert DJ, Sutherland GR, Jenkins NA, Herzog H: **GPR56, a novel secretin-like human G-protein-coupled receptor gene.** *Genomics* 1999, 55:296-305.
59. McLatchie LM, Fraser NJ, Main MJ, Wise A, Brown J, Thompson N, Solari R, Lee MG, Foord SM: **RAMPs regulate the transport and ligand specificity of the calcitonin- receptor-like receptor.** *Nature* 1998, 393:333-339.
60. Muff R, Born W, Fischer J: **Adrenomedullin and related peptides: receptors and accessory proteins.** *Peptides* 2001, 22:1765-1772.
61. **National Center for Biotechnology Information** [http://www.ncbi.nlm.nih.gov]
62. **The Swiss 7TM Search Tool** [http://www.expasy.ch/cgi-bin/search-7tm]
63. **European Bioinformatics Institute** [http://www.ebi.ac.uk]
64. **GPCRDB: Information system for G protein-coupled receptors (GPCRs)** [http://www.gpcr.org]
65. **PubMed** [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed]
66. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, 25:4876-4882.
67. Henikoff S, Henikoff JG: **Amino acid substitution matrices from protein blocks.** *Proc Natl Acad Sci USA* 1992, 89:10915-10919.
68. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp Ser* 1999:95-98.
69. **GeneDoc** [http://www.psc.edu/biomed/genedoc/]
70. Hofmann K, Stoffel W: **TMbase - A database of membrane spanning proteins segments.** *Biol Chem Hoppe-Seyler* 1993, 374:166.
71. **TMpred - Prediction of Transmembrane Regions and Orientation** [http://www.ch.embnet.org/software/TMPRED_form.html]
72. **PHYLIP** [http://evolution.genetics.washington.edu/phylip.html]
73. Jones DT, Taylor WR, Thornton JM: **The rapid generation of mutation data matrices from protein sequences.** *Comput Appl Biosci* 1992, 8:275-282.
74. Page RD: **TreeView: an application to display phylogenetic trees on personal computers.** *Comput Appl Biosci* 1996, 12:357-358.
75. Schmidt HA, Strimmer K, Vingron M, von Haeseler A: **TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing.** *Bioinformatics* 2002, 18:502-504

Trace amines: Identification of a family of mammalian G protein-coupled receptors

Beth Borowsky*, Nika Adham†, Kenneth A. Jones‡, Rita Raddatz, Roman Artymyshyn, Kristine L. Ogozalek, Margaret M. Durkin, Parul P. Lakhlani§, James A. Bonini, Sudam Pathirana, Noel Boyle, Xiaosui Pu, Evguenia Kouranova, Harvey Lichtblau, F. Yulina Ochoa, Theresa A. Branchek, and Christophe Gerald

Synaptic Pharmaceutical Corporation, Paramus, NJ 07652

Edited by L. L. Iversen, University of Oxford, United Kingdom, and approved May 17, 2001 (received for review March 2, 2001)

Tyramine, β-phenylethylamine, tryptamine, and octopamine are biogenic amines present in trace levels in mammalian nervous systems. Although some "trace amines" have clearly defined roles as neurotransmitters in invertebrates, the extent to which they function as true neurotransmitters in vertebrates has remained speculative. Using a degenerate PCR approach, we have identified 15 G protein-coupled receptors (GPCR) from human and rodent tissues. Together with the orphan receptor PNR, these receptors form a subfamily of rhodopsin GPCRs distinct from, but related to the classical biogenic amine receptors. We have demonstrated that two of these receptors bind and/or are activated by trace amines. The cloning of mammalian GPCRs for trace amines supports a role for trace amines as neurotransmitters in vertebrates. Three of the four human receptors from this family are present in the amygdala, possibly linking trace amine receptors to affective disorders. The identification of this family of receptors should rekindle the investigation of the roles of trace amines in mammalian nervous systems and may potentially lead to the development of novel therapeutics for a variety of indications.

Norepinephrine (NE), dopamine (DA), and serotonin (5-HT) are classical biogenic amine neurotransmitters whose well characterized effects are mediated by interactions with subfamilies of receptors that belong to the rhodopsin superfamily of G protein-coupled receptors (GPCRs). In addition to these classical amines, there exists a class of "trace amines" that are found in very low levels in mammalian tissues, and include tyramine, β-phenylethylamine (β-PEA), tryptamine, and octopamine (1). The rapid turnover of trace amines, as evidenced by their dramatic increases following treatment with monoamine oxidase (MAO) inhibitors or deletion of the MAO genes, suggests that the levels of trace amines at neuronal synapses may be considerably higher than predicted by steady-state measures (2–5). The role of trace amines as neurotransmitters in invertebrates is well established and octopamine is thought to be the sympathetic nervous system counterpart to NE (6–9). GPCRs for tyramine and octopamine have been cloned from both insects (10–14) and mollusks (15, 16).

Although there is clinical literature that supports a role for trace amines in depression as well as other psychiatric disorders and migraine (2, 3, 17–20), the role of trace amines as neurotransmitters in mammalian systems has not been thoroughly examined. Because they share common structures with the classical amines and can displace other amines from their storage vesicles, trace amines have been referred to as "false transmitters" (21). Thus, many of the effects of trace amines are indirect and are caused by the release of endogenous classical amines. However, there is a growing body of evidence suggesting that trace amines function independently of classical amine transmitters and mediate some of their effects via specific receptors (for review, see refs. 22–24). Saturable, high-affinity binding sites for [^{3}H]tryptamine (23, 25–27), *p*-[^{3}H]tyramine (28–30), and β-[^{3}H]PEA (31) have been reported in rat brain, and both the pharmacology and localization of these sites suggest that they are distinct from the amine transporters. However, although binding sites in brain and other tissues have been reported, no specific receptors for these trace amines have yet been identified conclusively.

We now report the identification of a family of related mammalian GPCRs of which two members have been shown to specifically bind and/or be activated by trace amines. TA_1 is activated most potently by tyramine and β-PEA, and TA_2 is activated most potently by β-PEA. The 15 distinct receptors described here, along with the orphan receptor PNR (32) and the pseudogenes GPR58, GPR57 (33), and the 5-HT_4 pseudogene (34), share a high degree of sequence homology and together form a subfamily of rhodopsin GPCRs distinct from but related to 5-HT, DA, and NE receptors. We further describe the localization of TA_1 in human and rodent tissues, as well as the chromosomal localization of the human members of this family. The identification of this family of receptors should facilitate the understanding of the roles of trace amines in the mammalian nervous system.

Materials and Methods

Degenerate PCR. To clone a rat TA_1 fragment, PCR was performed on genomic DNA by using primers designed based on an alignment of the sixth (5′-TNNKNTGYTGGYTNCCNTTYTTY-3′) and seventh (5′-ARNSWRTTNVNRTANCCNARCC-3′) transmembrane (TM) domains of a subset of 5-HT receptors. To clone rat TA_4, human TA_5, rat TA_7, rat TA_8, and rat TA_9, PCR was performed on genomic DNA by using primers designed based on an alignment of the first intracellular loop and TMII (5′-TTYAARCARYTNCAYWSNCCNAC-3′) or the first extracellular loop (5′-GARHVNTGYTGGTAYTTYGG-3′) and TMVI (5′-ATNCCNARNGTYTTNRCNGCYTT-3′ or 5′-CCARCANRNNARRAANACNCC-3′) of TA_1, GPR58, and GPR57. To clone TA_2, PCR was performed on rat genomic DNA by using primers designed based on an alignment of the first intracellular loop TMII (5′-TTYAARSMNYTNCAYWSNCCNAC-3′) and the first extracellular loop (5′-CCRAARWACCARCANBNYTCNRY-3′) of TA_3, TA_1, GPR58, PNR, and the 5-HT_4 pseudogene. For the cloning of a rat TA_3 fragment, PCR was performed on genomic DNA by using primers designed based on an alignment of TMVI (5′-GYNTWYRYNNTNWSNTGGHTNCC-3′) and TMVII (5′-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: 5-HT, serotonin; CFTR, cystic fibrosis transmembrane conductance regulator; CNS, central nervous system; DA, dopamine; GPCR, G protein-coupled receptor; MAO, monoamine oxidase; NE, norepinephrine; β-PEA, β-phenylethylamine; TA, trace amine; TM, transmembrane domain.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF380185–AF380203).

*To whom reprint requests should be addressed at: Synaptic Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652. E-mail: bborowsky@synapticcorp.com.

†Present address: Schering-Plough, Kenilworth, NJ 07033.

‡Present address: AstraZeneca R&D Boston, Waltham, MA 02451.

§Present address: University of Kentucky Medical Center, Lexington, KY 40536.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

www.pnas.org/cgi/doi/10.1073/pnas.151105198

AVNADNGBRWAVANNANNGGRTT-3') of a collection of rhodopsin GPCRs. PCR conditions were: 94°C for 5 min; 10 cycles of 94°C for 30 s, 44°C for 45 s or 43°C for 1 min (TA_1), 72°C for 1 min 45 s; 30 cycles of 94°C for 30 s, 49°C for 45 s or 48°C for 1 min (TA_1), 72°C for 1 min 45 s; 72°C for 20 min. PCR products were subcloned into the TA cloning vector (Invitrogen), sequenced (Big Dye cycle sequencing protocol and ABI 377 sequencers from Applied Biosystems), and analyzed (WISCONSIN Package, Genetics Computer Group, Madison, WI).

Library Screening. Rat liver or human placental genomic phage libraries (Stratagene) or a rat cosmid library (CLONTECH) were screened by using ^{32}P-labeled oligonucleotide probes and standard protocols. Positive signals were isolated and hybridizing bands, identified by Southern blot analysis, were subcloned into pcDNA3.1 (Invitrogen) or a modified form of pcEXV (35) and sequenced as above.

Low Stringency PCR. Fragments of species homologues of TA_1 were amplified from genomic DNA using primers designed against the rat TA_1. PCR was performed with the Expand Long Template PCR System (Roche Molecular Biochemicals) with an annealing temperature of 45–51°C.

Rapid Amplification of cDNA Ends (RACE). 5′ and 3′ RACE were performed according to the manufacturer's protocol, using Marathon-Ready cDNA (CLONTECH) from kidney and stomach (human TA_1), kidney and testes (rat TA_2), spinal cord (rat TA_3), and brain (mouse TA_1). Coding regions were amplified multiple times from genomic DNA, human amygdala cDNA, or rat testes cDNA by using primers specific to the 5′ and 3′ untranslated regions.

Oocyte Injection and Recording. Oocytes were isolated from *Xenopus laevis* (Xenopus 1, Ann Arbor, MI) and maintained, injected, incubated, and recorded from as described (36). Oocytes were injected with 10–15 ng of mRNA encoding TA_1 with or without 10 ng of mRNA encoding the cystic fibrosis transmembrane conductance regulator (CFTR; ref. 37). Ligands were applied by local perfusion from a 10-μl glass capillary tube 0.5 mm from the oocyte.

Measurement of Intracellular cAMP. Transiently transfected COS-7 cells were incubated in Dulbecco's PBS supplemented with 10 mM Hepes, 10 mM glucose, 5 mM theophylline, and 10 μM pargyline for 20 min at 37°C in 95% O_2/5% CO_2. Test compounds were added and cells were incubated for 10 min. The medium was aspirated and the reaction stopped by the addition of 200 μl of 100-mM HCl. The cAMP content in each well was measured by RIA (Scintillation Proximity Assay; Amersham Pharmacia Biotech) using a microbeta Trilux counter (Wallac, Gaithersburg, MD).

Radioligand Binding. Membranes prepared from cells transiently transfected with human TA_1 and rat $G\alpha_s$ were diluted in 25 mM Gly-Gly buffer (Sigma, pH 7.4 at 0°C) containing 5 mM ascorbate (final protein concentration = 120 μg/ml). Membranes were then incubated with [^3H]tyramine [American Radiochemicals, St. Louis; specific activity 60 mCi/μM (1 Ci = 37 GBq)] in the presence or absence of competing ligands on ice for 30 min in a volume of 250 μl. Bound ligand was separated from free ligand by filtration through GF/B filters presoaked in 0.5% polyethyleneimine, using a Brandel (Bethesda, MD) cell harvester vacuum filtration device, and bound radioactivity quantified by using a scintillation counter. Data were fit to nonlinear curves by using PRISM (GraphPad, San Diego).

Quantitative Reverse Transcription (RT)–PCR. cDNA was prepared from DNase-treated total RNA purchased from CLONTECH or isolated from human tissues by using TRIzol reagent (Life Technologies, Grand Island, NY). Integrity of RNA and cDNA was assessed by amplification of cyclophilin or glyceraldehyde-3-phosphate dehydrogenase (GAPDH). PCR reactions were carried out in a PE7700 sequence detection system (Perkin–Elmer) according to the manufacture's protocol. The probe [5′(6-FAM)-ATGGTGAGATCTGCTGAGCACTGTTGG-TATT-(TAMRA)3′] was labeled with FAM (6-carboxyfluorescein) as the reporter and TAMRA (6-carboxy-4,7,2,7′-tetramethylrhodamine) as a quencher, and the forward and reverse PCR primers were 5′-CATGGCCACTGTGGACTT-TCT-3′ and 5′-GTCGGTGCTTGTGTGAATTTTACA-3′, respectively. The fluorescent signal from each well was normalized by using an internal passive reference, and data were fitted to a standard curve generated with genomic DNA.

Chromosomal Localization. The Stanford Human Genome Center (SHGC) G3 panel of 83 radiation hybrids was analyzed by PCR using 20 ng of DNA and the same primers, probes, and thermal cycler profiles as used for localization. The RH Server (at SHGC: www-SHGC.stanford.edu) and the National Center for Biotechnology Information's LocusLink and GeneMap '99 were used for analysis.

***In Situ* Hybridization Histochemistry.** Sense and antisense riboprobes (251 bp, TMV-TMVI of mouse TA_1) were labeled with digoxigenin as outlined in the DIG/Genius System (Roche Molecular Biochemicals). Male 129S6/SVEV mice (20 g, Taconic Farms) were anesthetized with ketamine 20 mg/kg (Research Biochemicals) and xylazine 0.2 mg/kg (Sigma), and perfused transcardially with PBS followed by 4% paraformaldehyde/PBS. Tissues were cryoprotected, stored at −20°C, and sectioned (30 μM) by using a freezing microtome. Free-floating sections were incubated in 100 mM glycine for 5 min and 0.3% Triton X-100 for 15 min, then rinsed twice in PBS for 5 min. *In situ* hybridization histochemistry was carried out on free-floating tissue sections as outlined in the DIG/Genius System with a hybridization temperature of 52°C in a buffer containing 40% formamide.

Results

In an attempt to identify additional 5-HT_1-like receptors, such as the elusive 5-HT1p receptor (38), degenerate PCR primers were designed against TMs VI and VII of an alignment of 5-HT_1 receptors and used to amplify rat genomic DNA at reduced stringency. One product from this reaction was found to be a DNA sequence, not found in the GenBank database, with 42–48% amino acid identity to 5-HT_4, DA D_2, and β-adrenergic receptors. Sequencing of the corresponding full-length cDNA, BO111, revealed an ORF of 996 bp that is predicted to encode a protein of 332 aa (Fig. 1, rat TA_1). An allelic variant of this receptor was also identified wherein a glutamine replaces a leucine at position 170. BO111 is most closely related to GPR58 (50% aa identity), the human 5-HT_4 pseudogene (47% aa identity, with frame shifts "corrected"), BO107 (an orphan GPCR previously identified at Synaptic and later renamed TA_3) and GPR57 (45% aa identities), PNR (38% aa identity), and 5-HT_{1D}, 5-HT_4, and 5-HT_7 receptors (35–37% aa identities). Human and mouse orthologues of BO111 were obtained by standard methods. The amino acid sequences of the human and mouse receptors are 76% identical to each other and 79% and 87% identical to the rat receptor, respectively (Fig. 1).

A search for the endogenous ligand for the receptor encoded by BO111 was performed by expressing it in oocytes along with mRNA encoding the cAMP-responsive Cl channel, CFTR. Candidate ligands were tested in eleven groups of five. From this broad panel, octopamine and, more weakly, DA and 5-HT, elicited inward currents at 100 μM (Fig. 2*A*). Stimulation by

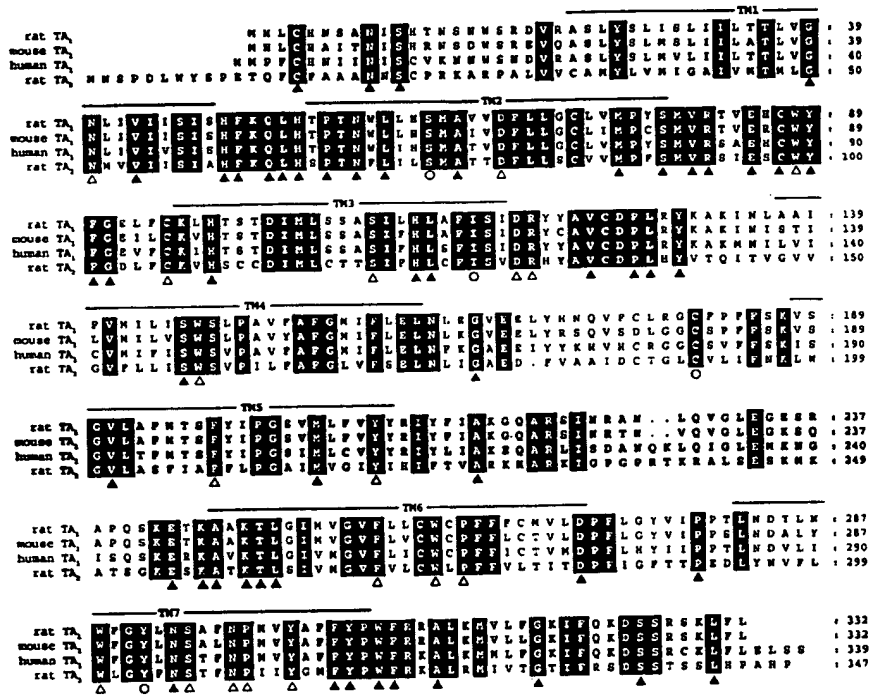


Fig. 1. Alignment of rat, mouse, and human TA_1 and rat TA_2 receptors (GenBank accession nos. AF380186, AF380187, AF380185, and AF380188, respectively). Shaded residues are conserved in all four receptors. Triangles and circles indicate residues conserved in TA_1–TA_{15}. Open triangles are also conserved among all human monaminergic receptors, and open circles are conserved among all human 5-HT but not NE or DA receptors. Seven putative TM domains are indicated.

octopamine (100 μM) produced an average current amplitude of 230 ± 55 nA (n = 4). Similar currents were generated by tyramine at a lower concentration (100 nM; 287 ± 31 nA, n = 28; Fig. 2*B*). EC_{50} values were obtained for octopamine (635 ± 151 nM) and tyramine (37 ± 4.4 nM) from cumulative concentration effect responses (data not shown). These results suggested that BO111 encodes a receptor for trace amines, and was thus named TA_1. No such currents were observed in oocytes injected with only mRNA encoding the CFTR channel. Oocytes expressing rat TA_1 without CFTR failed to generate inward currents (Fig. 2*C*; n = 11), suggesting that stimulation of rat TA_1 by octopamine and tyramine resulted in the generation of cAMP leading to CFTR channel opening, presumably via activation of the endogenous *Xenopus* G protein $G\alpha_s$. Oocytes expressing the human orthologue of rat TA_1 with CFTR also produced inward currents in response to application of 100 nM tyramine (Fig. 2*D*).

Additional bioamines were tested for activity at human TA_1 expressed in mammalian cells. Human TA_1 was activated most potently by β-PEA and tyramine, and more weakly by octopamine and DA (Table 1 and Fig. 3). The agonists listed in Table 1 produced an increase in intracellular cAMP accumulation, likely via the $G\alpha_s$-class of G proteins in COS-7 cells transfected with human TA_1, but not in mock-transfected cells.

Consistent with the relatively high potency of tyramine for activating human TA_1, [^{3}H]tyramine demonstrated high-affinity, saturable binding in TA_1-expressing membranes (average K_d = 20 nM; data not shown). Selectivity of human TA_1 for β-PEA

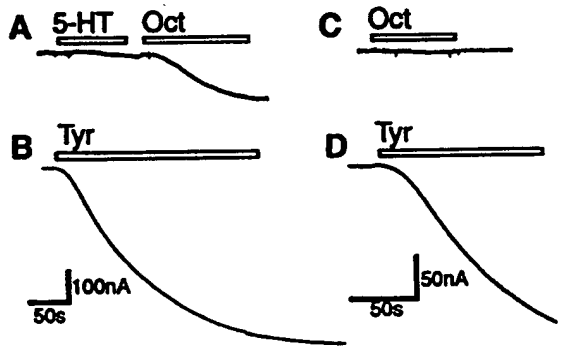


Fig. 2. Voltage-clamp responses to trace amines in oocytes. (*A*) Response to 100 μM octopamine (Oct) or 5-HT in an oocyte expressing rat TA_1 and CFTR. (*B*) Response to 100 nM tyramine (Tyr) in an oocyte expressing rat TA_1 and CFTR. (*C*) Response to 100 μM octopamine in an oocyte expressing only TA_1. (*D*) Response to 100 nM tyramine (Tyr) in an oocyte expressing human TA_1 and CFTR. Holding potential was −80 mV for all oocytes. Marker bar in *D* also applies to *A* and *C*.

Table 1. Pharmacological profile of human TA_1

Compound	K_i, nM	EC_{50}, nM
β-PEA	8.0 ± 3.2	324 ± 110
Tyramine	34 ± 11	214 ± 67
Dopamine	422 ± 11	6,700 ± 1,700
Octopamine	493 ± 99	4,029 ± 75
Tryptamine	1,084 ± 159	>6 uM
Histamine	3,107 ± 1,593	>5 uM
Serotonin	>6 uM	>10 uM
Norepinephrine	>10 uM	>5 uM

Values represent the average ± SEM from $n \geq 3$ experiments. Binding K_is were determined from displacement of [^{3}H]tyramine (20 nM) and EC_{50} values were determined by increases in cAMP accumulation.

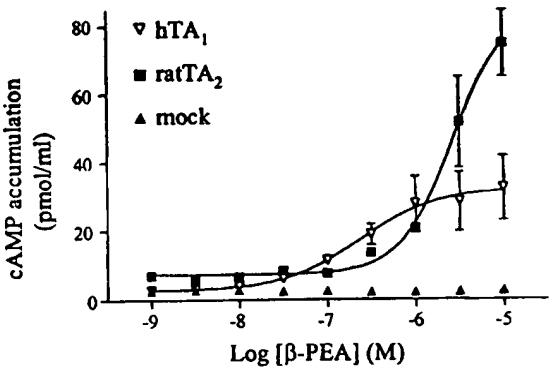


Fig. 3. β-PEA-induced responses in COS-7 cells transfected with human TA_1, rat TA_2, or vector. Cells were incubated with increasing concentrations of β-PEA and cAMP accumulation measured. Data are from duplicate determinations and are representative of three to six experiments.

and tyramine is shown in Table 1. The rank order of potency was similar between the binding and functional assays, although the K_i values determined from binding displacement were ≈6-fold lower than EC_{50} values determined in functional studies. This difference has been reported for many exogenously expressed receptors, including the 5-HT_7 receptor (39), and may be due to the relatively low expression levels of TA_1 in COS-7 cells, weak coupling of the receptor to signaling components in these cells, or differences in assay conditions.

Additional degenerate PCR work performed on rat genomic DNA led to the identification of TA_2. TA_2 is most similar to the human 5-HT_4 pseudogene (82% aa identity with frame shifts "corrected"), and shares 48–51% aa identity to the rat and human TA_1 receptors (Fig. 1), GPR57 and GPR58. The expression of rat TA_2 in COS-7 cells resulted in an increase in cAMP accumulation, presumably via $G\alpha_s$-class G protein(s). Of the biogenic amines tested, only β-PEA and tryptamine activated this receptor; however, the response was of low potency [EC_{50} = 1.9 ± 0.5 μM (Fig. 3) and 17 ± 2 μM (data not shown) respectively]. The low potency of trace amines for rat TA_2 in heterologous expression systems may be explained by its poor surface expression, as determined by subcellular localization of an epitope-tagged rat TA_2 (data not shown). Alternatively, other more potent agonists may exist for rat TA_2. Because the human orthologue of this receptor is most likely the 5-HT_4 pseudogene, no further studies were conducted on rat TA_2.

Further degenerate PCR work led to the identification of TA_5 from human genomic DNA, and TA_4, TA_7, TA_8, and TA_9 from rat genomic DNA. While isolating these full-length receptors from genomic libraries, several additional closely related receptors were also isolated, including human TA_4 and rat TA_6, TA_{10}, TA_{11}, TA_{12}, TA_{13}, TA_{14}, and TA_{15}. TA_4–TA_{15} are highly homologous to each other, with overall aa identities of 62–96% (see Fig. 6, which is published as supplemental data on the PNAS web site, www.pnas.org). These receptors are 66–73% identical to TA_3, 41–48% identical to TA_2, 40–44% identical to TA_1, and 28–36% identical to 5-HT receptors. As indicated in Fig. 1, there are 74 residues that are completely conserved in TA_1–TA_{15}. Of these, 52 are uniquely conserved in the trace amine family (closed triangles), 18 (of 25) are also conserved in all human monoaminergic receptors (open triangles), and an additional 4 (of 32) are conserved with human 5-HT receptors, but not NE or DA receptors (open circles).

A phylogenetic tree was constructed from the aa sequences of TA_1–TA_{15}, PNR, GPR57, GPR58, the 5-HT_4 pseudogene, and several vertebrate and invertebrate aminergic receptors (Fig. 4).

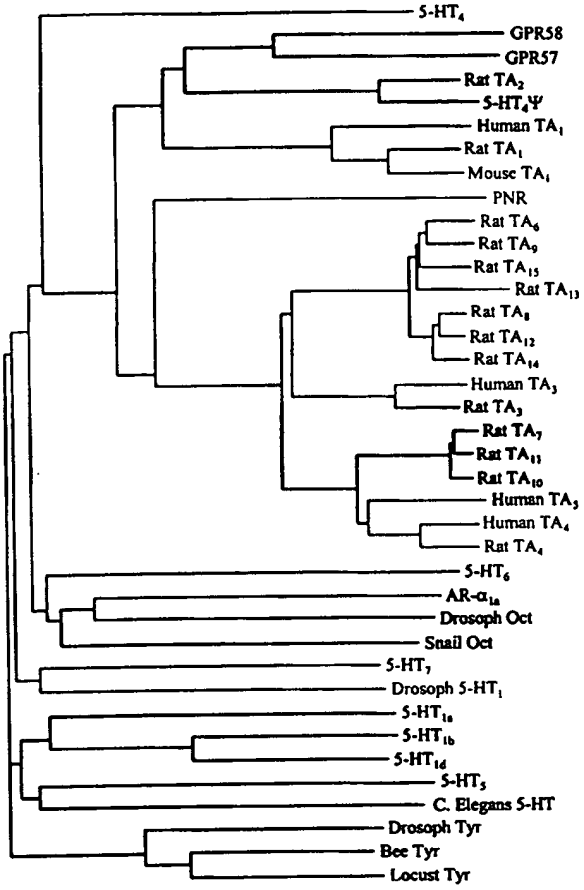


Fig. 4. A phylogenetic tree for trace amine receptors TA_1–TA_{15}, human 5-HT receptors 5-HT_{1a}, 5-HT_{1b}, 5-HT_{1d}, 5-HT_4, 5-HT_5, 5-HT_7, human α1a receptor (AR-α_{1a}), GPR57, GPR58, PNR, 5-HT_4 psuedogene (5-$HT_{4\psi}$), drosophila (Drosoph) receptors for octopamine (Oct), 5-HT_1 and tyramine (Tyr), *Caenorhabditis elegans* (C. Elegans) 5-HT receptor, tyramine receptors from bee (Bee Tyr) and locust (Locust Tyr), and a snail octopamine receptor (Snail Oct). Amino acid sequences for each receptor spanning from the start of TMI to the end of TMVII were aligned by using the CLUSTALW algorithm and the tree constructed by the NJ method on a Decypherll Bioaccelerator (TimeLogic, Reno, NV). GenBank accession nos.: AF380190 (rat TA_3), AF380189 (human TA_3), AF380191 (rat TA_4), AF380192 (human TA_4), AF380193-AF380203 (TA_5–TA_{15}).

The TA receptors, along with GPR57, GPR58, 5-HT_4 pseudogene, and PNR, branch separately from mammalian receptors for classical biogenic amines, including those for 5-HT, and from the invertebrate trace amine receptors. Within this large family of receptors, there appears to be at least two subfamilies. TA_1 and TA_2, along with GPR57, GPR58, and 5-HT_4 pseudogene, constitute one subfamily, and TA_3–TA_{15} constitute a second subfamily.

Radiation hybrid mapping using primers selective for TA_1, TA_3, TA_4, and TA_5 was used to identify the chromosomal localization of the human trace amine receptors. All four genes had virtually identical patterns and mapped to SHGC-1836. This placed the TA receptor genes in the region of chromosome 6q23.2. Interestingly, PNR, GPR57, GPR58, and 5-HT_4 pseudogene were previously shown to be clustered between 6q22 and 6q24 by using fluorescence *in situ* hybridization analysis (32, 33).

Human TA_1 mRNA was detected by quantitative reverse transcription (RT)–PCR in low levels in discrete regions within the central nervous system (CNS) and in several peripheral

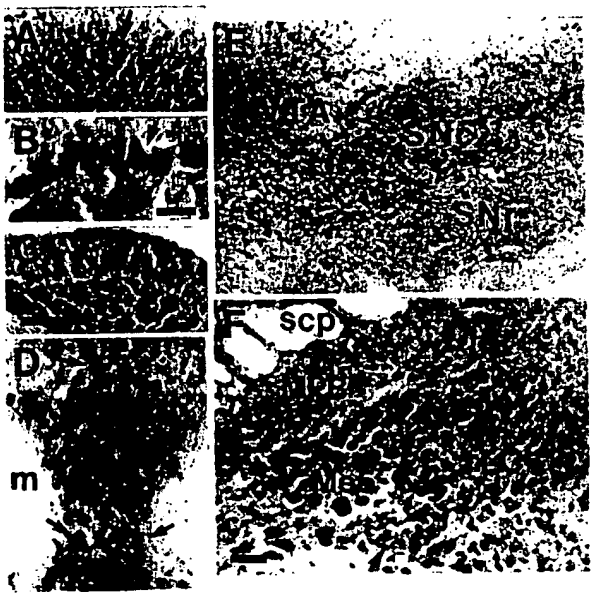


Fig. 5. Photomicrographs showing hybridization signals for TA_1 mRNA in mouse CNS. Signal detected in cerebellar Purkinje cells (arrows) hybridized with antisense (*A* and *B*) and sense (*C*) probes. Scale bar in *C* (200 μm) also applies to *A* and *E*; scale bar in *B*, 25 μm. Photomicrographs showing hybridization signal in cells (arrows) in the dorsal raphe (*D*), ventral tegmental area (VTA), substantia nigra, compact part (SNc) and reticular part (SNr) (*E*), and locus coeruleus (LC; *F*). Scale bar in *F* (50 μm) also applies to *D*. MPB, medial parabrachial nucleus; Me5, mesencephalic trigeminal nucleus; m, medial longitudinal fasciculus; scp, superior cerebral peduncle.

tissues. Moderate levels (100 copies/ng cDNA) were expressed in stomach, low levels (15–100) expressed in amygdala, kidney, lung, and small intestine, whereas trace amounts (<15) were expressed in cerebellum, dorsal root ganglia, hippocampus, hypothalamus, liver, medulla, pancreas, pituitary, pontine reticular formation, prostate, skeletal muscle, and spleen. Message levels for the other human trace amine receptors were also detected in low levels. TA_3 mRNA was detected only in kidney. TA_4 and TA_5 mRNA were expressed in kidney and amygdala, and TA_4 was also detected in the hippocampus.

A widespread and unique distribution of TA_1 mRNA was revealed in the mouse CNS by *in situ* hybridization histochemistry (Fig. 5) and the hybridization signal was localized to the cytoplasm of neuronal profiles (Fig. 5 *A* and *B*). Several brain regions exhibited intense labeling specifically, the mitral cell layer of the olfactory bulb, piriform cortex, the arcuate, motor, and mesencephalic trigeminal nuclei, lateral reticular and hypoglossal nuclei, cerebellar Purkinje cells, and ventral horn of the spinal cord. Moderate labeling was evident in the frontal, entorhinal, and agranular cortices, the ventral pallidum, thalamus, hippocampus, several hypothalamic nuclei, ambiguus, dorsal raphe, and gigantocellular reticular nuclei. Weaker staining was visible in the septum, basal ganglia, amygdala, myelencephalon, and spinal cord dorsal horn. Particularly interesting was the moderate expression of TA_1 mRNA in several monoaminergic cell groups, namely the dorsal raphe (Fig. 5*D*), the locus coeruleus (Fig. 5*F*), and the ventral tegmental area (Fig. 5*E*).

Discussion

We have identified a multigene family of intronless GPCRs and have demonstrated that TA_1 is potently activated by tyramine and β-PEA and displays low affinity for tryptamine, octopamine, and DA. An additional member of this family, TA_2, is also activated by β-PEA and tryptamine. Although the roles of tyramine and octopamine as neurotransmitters acting via stimulation of G protein-coupled receptors in invertebrate systems are well established (6–9), mammalian GPCRs for trace amines have not, to our knowledge, been previously reported. The present finding lends strong support to a role for trace amines as neurotransmitters or neuromodulators in vertebrates.

One of the most interesting and unexpected findings of this study was the discovery of such a large family of highly related receptors. In addition to TA_1 and TA_2, we have identified 13 other related receptors. These additional receptors, TA_3–TA_{15}, share an unusually high degree of amino acid identity (62–96%). For comparison, the human 5-HT receptors share 28–63% amino acid identities. The high degree of homology within members of the TA family and the tight clustering of the human TA receptors on chromosome 6q23.2 suggests that these receptors evolved relatively recently, after the invertebrate/vertebrate split, and makes it tempting to speculate that this region may represent a hotspot for gene duplication events.

Although the degree of homology between receptors within a species is extremely high, the degree of amino acid identity among orthologues is moderate to low. The rat and human orthologues of TA_3 and TA_4 share a moderate degree of amino acid identities (87% and 88%, respectively). However, the mouse, rat, and human orthologues of TA_1 share a relatively low degree of homology (87% for rat and mouse, 79% for rat and human, 76% mouse and human). This observation suggests that although these receptors are relatively recent expansions of the genome, they are evolving at a rapid rate.

Another interesting observation is that a larger number of rat receptors has been identified so far as compared with human receptors. Four human receptors have been identified (TA_1, TA_3, TA_4, and TA_5), whereas 14 rat receptors have been identified (TA_1–TA_4 and TA_6–TA_{15}). To date, only a human form of PNR has been reported (29). There are also a large number of pseudogenes within the human members of this family. The 5-HT_4 pseudogene (31), which we propose should be renamed ψTA_2, and ψGPR57 (33) each contain frame shifts resulting in premature stop codons, whereas ψGPR58 lacks an amino terminus (33). Although additional human receptors may ultimately be identified, the striking difference in the number of rat and human receptors suggests that this family may play very different roles in different species.

We have demonstrated a functional response to heterologously expressed TA_1 in both *Xenopus* oocytes and a mammalian cell system. The response in both assays indicates that TA_1 couples to the stimulation of adenylate cyclase through a $G\alpha_s$ G protein. The human TA_1 receptor is activated by tyramine and β-PEA, less potently by octopamine, and binds β-PEA and tyramine with high affinity and tryptamine, octopamine, and DA with lower affinity. The rat TA_2 receptor is activated by β-PEA and tryptamine, also via stimulation of a $G\alpha_s$ G protein. Thus far, we have not demonstrated functional responses to tyramine, β-PEA, tryptamine, octopamine, or the classical biogenic amines in COS-7 cells expressing TA_3–TA_{15}. This finding may be due to poor trafficking to the plasma membrane (data not shown), or these receptors may respond to related, perhaps as yet unidentified, amines. However, the high degree of sequence conservation between the two subfamilies, the evolutionary branching analysis, as well as the chromosomal proximity of the receptors make it very likely that TA_3–TA_{15} encode receptors for trace amines.

Human TA_1 mRNA is expressed in low to moderate levels in peripheral tissues such as stomach, kidney, and lung, and within the CNS appears to be restricted primarily to the amygdala. The expression of TA_1 mRNA is lower than that seen for receptors of classical neurotransmitters, consistent with the low levels of trace amines relative to other neurotransmitters. All of the human members of the TA family are expressed in the kidney, supporting

a role in blood pressure regulation and electrolyte homeostasis. This may be related to the "cheese effect," wherein dietary-induced elevations in tyramine levels in patients taking MAO inhibitors results in hypertension and migraine (see ref. 19 for review).

The expression of TA_1 mRNA in human amygdala is intriguing in light of evidence suggesting a role of trace amines in the etiology and/or treatment of depression and anxiety disorders. A functional deficiency of β-PEA and tryptamine has been proposed as a potential etiological factor in depression (40–42), and increased levels of β-PEA are associated with the manic phase of bipolar disease (43). Antidepressants that inhibit MAO produce proportionally greater increases in trace amines than 5-HT (2, 3). MAO-B knockout mice have increased levels of β-PEA (7-fold higher), but normal levels of 5-HT, NE, and DA (5). Interestingly, MAO-B knockout mice show a reduced decrease in mobility in the forced swim test, similar to that induced by antidepressants (5). Taken together, these results suggest that TA_1 receptors in the amygdala may be an important site of action for trace amines, particularly β-PEA, in the etiology and treatment of depression. The expression of mouse TA_1 mRNA in the dorsal raphe, locus ceruleus, and ventral tegmental area indicates that trace amines may modulate the activity of 5-HT, NE, and DA systems and further supports a role for trace amine receptors in the regulation of mood.

Human trace amine receptor genes map to chromosome 6q23.2, close to SCZD5, a susceptibility locus for schizophrenia (6q13-26 with the greatest allele sharing at 6q21-22.3; ref. 44). Because of structural and physiological similarities, β-PEA has been described as the body's endogenous amphetamine (45, 46). Amphetamine produces a paranoid schizophrenic syndrome in humans, and chronic treatment with either amphetamine or β-PEA produces a behavioral sensitization in animals (47–51). Moreover, numerous clinical studies have demonstrated elevated urinary levels of β-PEA in schizophrenic patients (52, 53). Thus, it will be important to delineate the role of TA receptors in the etiology and treatment of schizophrenia.

Although trace amines have long been thought to be neurotransmitters, the understanding of their physiology has lagged that of the classical biogenic amines, in part, because the receptor targets remained elusive. The identification of mammalian GPCRs for trace amines supports a role for trace amines as bona fide neurotransmitters in vertebrates. The localization of mRNA for three of the four human receptors in amygdala lends a potential site of action for the postulated role of trace amines in the etiology and/or treatment of several affective disorders. Future characterization of TA_3–TA_{15} will further enhance our understanding of these receptors. The discovery of this family of receptors provides a means to evaluate the physiological roles of trace amines in higher species and their regulation in diseased processes, and to explore potential therapeutic applications associated with these receptors.

The authors thank Tracy Johnson-Blake and Stacy Kokkinakis for cell culture, Debby Tambe for plasmid preparation, Siqun Zhou and Meng Dai for excellent technical help, Dan Larhammar for insight on phylogenic analysis, George Moralishvili for assistance in manuscript preparation, and Elisabeth Griggs for photographic assistance.

1. Usdin, E. & Sandler, M., eds. (1976) *Trace Amines and the Brain* (Dekker, New York).
2. Boulton, A. A. (1976) in *Trace Amines and the Brain*, eds. Usdin, E. & Sandler, M. (Dekker, New York).
3. Juorio, A. V. (1976) *Brain Res.* **111**, 442–445.
4. Durden, D. A. & Philips, S. R. (1980) *J. Neurochem.* **34**, 1725–1732.
5. Grimsby, J., Toth, M., Chen, K., Kumazawa, T., Klaidman, L., Adams, J. D., Karoum, F., Gal, J. & Shih, J. C. (1997) *Nat. Genet.* **17**, 206–210.
6. Axelrod, J. & Saavedra, J. M. (1977) *Nature (London)* **265**, 501–504.
7. David, J. C. & Coulon, J.-F. (1985) *Prog. Neurobiol.* **24**, 141–185.
8. Evans, P. D. & Robb, S. (1993) *Neurochem. Res.* **18**, 869–874.
9. Roeder, T. (1999) *Prog. Neurobiol.* **59**, 533–531.
10. Arakawa, S., Gocayne, J. D., McCombie, W. R., Urquhart, D. A., Hall, L. M., Fraser, C. M. & Venter, J. C. (1990) *Neuron* **2**, 343–354.
11. Saudou, F., Amlaiky, N., Plassat, J.-L., Borelli, E. & Hen, R. (1990) *EMBO J.* **9**, 3611–3617.
12. Vanden Broeck, J., Vulsteke, V., Huybrechts, R. & DeLoof, A. (1995) *J. Neurochem.* **64**, 2387–2395.
13. Han, K.-A., Millar, N. S. & Davis, R. L. (1998) *J. Neurosci.* **18**, 3650–3658.
14. Blenau, W., Balfanz, S. & Baumann, A. (2000) *J. Neurochem.* **74**, 900–908.
15. Gerhardt, C., Bakker, R. A., Piek, G. J., Planta, R. J., Vreugdenhil, E., Leysen, J. E. & Van Heerikhuizen, H. (1997) *Mol. Pharmacol.* **51**, 293–300.
16. Gerhardt, C., Lodder, H. C., Vincent, M., Bakker, R. A., Planta, R. J., Vreugdenhil, E., Kits, K. S. & Van Heerikhuizen, H. (1997b) *J. Biol. Chem.* **272**, 6201–6207.
17. Coppen, A., Shaw, D. M., Malleson, A., Eccleston, E. & Grundy, G. (1965) *Br. J. Psychiatry* **111**, 993–998.
18. Sandler, M., Ruthven, C. J. R., Goodwin, B. L., Reynolds, G. P., Rao, V. A. R. & Coppen, A. (1979) *Nature (London)* **278**, 357–358.
19. Vaughan, T. R. (1994) *Clin. Rev. Allergy* **12**, 167–180.
20. Merikangas, K. R., Stevens, D. E., Merikangas, J. R., Katz, C. B., Glover, V., Cooper, T. & Sandler, M. (1995) *Biol. Psychiatry* **38**, 730–736.
21. McGeer, P. L., Eccles, Sir J. C. & McGeer, E. G., eds. (1979) *Molecular Neurobiology of the Mammalian Brain* (Plenum, New York), p. 362.
22. Jones, R. S. G. (1982) *Prog. Neurobiol.* **19**, 117–139.
23. Perry, D. C. (1986) *J. Pharmacol. Exp. Ther.* **236**, 548–559.
24. Altar, C. A., Wasley, A. M. & Martin, L. L. (1986) *Neuroscience* **17**, 263–273.
25. Rudling, J. E., Richardson, J. & Evans, P. D. (2000) *Br. J. Pharmacol.* **131**, 933–941.
26. Kellar, K. J. & Cascio, C. S. (1982) *Eur. J. Pharmacol.* **78**, 475–478.
27. McCormack, J. K., Bietz, A. J. & Larson, A. A. (1986) *J. Neurosci.* **6**, 94–101.
28. Ungar, F., Mosnaim, A. D., Ungar, B. & Wolf, M. E. (1977) *Biol. Psychiatry* **12**, 661–668.
29. Vaccari, A. (1986) *Br. J. Pharmacol.* **89**, 15–25.
30. Vaccari, A. (1988) in *Trace Amines: Comparative and Clinical Neurobiology*, eds. Boulton, A. A., Juorio, A. V. & Downer, R. G. H. (Humana, Clifton, New Jersey), pp. 119–132.
31. Hauger, R. L., Skolnick, P. & Paul, S. M. (1982) *Eur. J. Pharmacol.* **83**, 147–148.
32. Zeng, Z., Fan, P., Rand, E., Kyaw, H., Su, K., Madike, K. C. & Li, Y. (1998) *Biochem. Biophys. Res. Comm.* **242**, 575–578.
33. Lee, D. K., Ltnch, K. R., Nguyen, T., Im, D.-S., Cheng, R., Salidiva, V. R., Liu, Y., Liu, I. S. C., Heng, H. H. Q., Seeman, P., *et al.* (2000) *Biochim. Biophys. Acta* **1490**, 311–323.
34. Liu, I. S. C., Kusumi, I., Ulpian, C., Tallerico, T. & Seeman, P. (1998) *Mol. Brain Res.* **53**, 98–103.
35. Miller, J. & Germain, R. N. (1986) *J. Exp. Med.* **164**, 1478–1489.
36. Smith, K. E., Forray, C., Walker, M. W., Jones, K. A., Tamm, J. A., Bard, J., Branchek, T. A., Linemeyer, D. L. & Gerald, C. (1997) *J. Biol. Chem.* **272**, 24612–24616.
37. Riordan, J. R. (1993) *Annu. Rev. Physiol.* **55**, 609–630.
38. Branchek, T. A., Mawe, G. M. & Gershon, M. D. (1988) *J. Neurosci.* **8**, 2582–2595.
39. Bard, J. A., Zgombick, J., Adham, N., Vaysse, P., Branchek, T. A. & Weinshank, R. L. (1993) *J. Biol. Chem.* **268**, 23422–23426.
40. Dewhurst, W. G. (1968) *Nature (London)* **218**, 1130–1133.
41. Dewhurst, W. G. & Marley, E. (1965) *Br. J. Pharmacol.* **25**, 705–727.
42. Sabelli, H. C. & Mosnaim, A. D. (1974) *Am. Psychiatry* **131**, 695–699.
43. Linnoila, M., Karoum, F., Cutler, N. R. & Potter, W. Z. (1983) *Biol. Psychiatry* **18**, 513–516.
44. Cao, Q., Martinez, M., Zhang, J., Sanders, A. R., Badner, J. A., Cravchik, A., Markey, C. J., Beshah, E., Guroff, J. J., Maxwell, M. E., *et al.* (1997) *Genomics* **43**, 1–8.
45. Sandler, M. & Reynolds, G. P. (1976) *Lancet* **1**, 70–71.
46. Wyatt, R. J. (1978) in *The Nature of Schizophrenia*, eds. Wynne, L. C., Cromwell, R. L. & Matthysse, S. (Wiley, New York), pp. 116–125.
47. Ellinwood, E. H., Sudilovsky, A. & Nelson, L. M. (1972) *Biol. Psychiatry* **4**, 215–225.
48. Eichler, A. J., Antelman, S. M. & Black, C. (1980) *Psychopharmacology* **68**, 287–290.
49. Randrup, A. & Munkvad, I. (1967) *Psychopharmacologia* **1**, 300–310.
50. Segal, D. S. & Janowsky, D. S. (1978) in *Psychopharmacology: A Generation of Progress*, eds. Lipton, M. A., DiMascio, A. & Killam, K. F. (Raven, New York), pp. 1113–1123.
51. Borison, R. L. & Diamone, B. I. (1978) *Biol. Psychiatry* **13**, 217–225.
52. O'Reilly, R. L. & Davis, B. A. (1994) *Prog. Neuropsychopharmacol. Biol. Psychiatry* **18**, 63–75.
53. Wolf, M. E. & Mosnaim, A. D. (1983) *Gen. Pharmacol.* **14**, 385–390.

ELSEVIER

Gene 275 (2001) 83–91

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Discovery and mapping of ten novel G protein-coupled receptor genes

Dennis K. Lee[a], Tuan Nguyen[b], Kevin R. Lynch[c], Regina Cheng[b], William B. Vanti[a], Oxana Arkhitko[a], Tressa Lewis[a], Jilly F. Evans[d], Susan R. George[a,b,e], Brian F. O'Dowd[a,b,*]

[a]*Department of Pharmacology, University of Toronto, Toronto, Ontario, M5S 1A8, Canada*
[b]*The Centre for Addiction and Mental Health, Toronto, Ontario, M5S 2S1, Canada*
[c]*Department of Pharmacology, University of Virginia Health Sciences Center, Charlottesville, VA 22908, USA*
[d]*Department of Pharmacology, Merck & Co. Inc., 770 Sumneytown Pike, P.O. Box 4, West Point, PA 19486, USA*
[e]*Department of Medicine, University of Toronto, Toronto, Ontario, M5S 1A8, Canada*

Received 10 May 2001; accepted 1 August 2001
Received by W. Makalowski

Abstract

We report the identification, cloning and tissue distributions of ten novel human genes encoding G protein-coupled receptors (GPCRs) *GPR78*, *GPR80*, *GPR81*, *GPR82*, *GPR93*, *GPR94*, *GPR95*, *GPR101*, *GPR102*, *GPR103* and a pseudogene, *ψGPR79*. Each novel orphan GPCR (oGPCR) gene was discovered using customized searches of the GenBank high-throughput genomic sequences database with previously known GPCR-encoding sequences. The expressed genes can now be used in assays to determine endogenous and pharmacological ligands. *GPR78* shared highest identity with the oGPCR gene *GPR26* (56% identity in the transmembrane (TM) regions). *ψGPR79* shared highest sequence identity with the $P2Y_2$ gene and contained a frame-shift truncating the encoded receptor in TM5, demonstrating a pseudogene. *GPR80* shared highest identity with the $P2Y_1$ gene (45% in the TM regions), while *GPR81*, *GPR82* and *GPR93* shared TM identities with the oGPCR genes *HM74* (70%), *GPR17* (30%) and $P2Y_5$ (40%), respectively. Two other novel GPCR genes, *GPR94* and *GPR95*, encoded a subfamily with the genes encoding the UDP-glucose and $P2Y_{12}$ receptors (sharing >50% identities in the TM regions). *GPR101* demonstrated only distant identities with other GPCR genes and *GPR102* shared identities with *GPR57*, *GPR58* and *PNR* (35–42% in the TM regions). *GPR103* shared identities with the neuropeptide FF 2, neuropeptide Y2 and galanin GalR1 receptors (34–38% in the TM regions). Northern analyses revealed *GPR78* mRNA expression in the pituitary and placenta and *GPR81* expression in the pituitary. A search of the GenBank databases with the *GPR82* sequence retrieved an identical sequence in an expressed sequence tag (EST) partially encoding *GPR82* from human colonic tissue. The *GPR93* sequence retrieved an identical, human EST sequence from human primary tonsil B-cells and an EST partially encoding mouse *GPR93* from small intestinal tissue. *GPR94* was expressed in the frontal cortex, caudate putamen and thalamus of brain while *GPR95* was expressed in the human prostate and rat stomach and fetal tissues. *GPR101* revealed mRNA transcripts in caudate putamen and hypothalamus. *GPR103* mRNA signals were detected in the cortex, pituitary, thalamus, hypothalamus, basal forebrain, midbrain and pons.

Keywords: Orphan G protein-coupled receptor; Transmembrane; Intronless; Pseudogene; Chromosome

1. Introduction

As is frequently stated, GPCRs are the largest family of cell surface receptors and are responsible for the signal transduction for a diverse variety of ligands including nucleotides, biogenic amines, peptides and other small molecules (Marchese et al., 1999). GPCRs share a common heptahelical topography and these regions are embedded in the membrane. These seven transmembrane (TM) regions share the most significant levels of receptor identity. As a consequence, the majority of DNA sequences encoding GPCRs were found using methods dependent on sequence homology, mainly PCR or electronic sequence database screening (Marchese et al., 1998). GPCRs activated by similar ligands share the greatest identities with each other.

Abbreviations: aa, amino acid; BAC, bacterial artificial chromosome; EST, expressed sequence tag; GPCRs, G protein-coupled receptors; HTGS, high-throughput genomic sequences; nr, non-redundant; oGPCRs, orphan G protein-coupled receptors; ORF, open reading frame; TM, transmembrane

* Corresponding author. Department of Pharmacology, Medical Science Building, Room 4353, University of Toronto, 8 Taddle Creek Road, Toronto, Ontario, Canada M5S 1A8. Tel.: +1-416-978-7579; fax: +1-416-978-2733.
E-mail address: brian.odowd@utoronto.ca (B.F. O'Dowd).

0378-1119/01/$ - see front matter
PII: S0378-1119(01)00651-5

However, newly discovered GPCRs frequently have only distant identities with known GPCRs, and these oGPCRs are difficult to characterize given the diversity in structure of the ligands and effector systems. This problem is compounded as we now realize that many endogenous ligands remain to be discovered. Increasingly, oGPCR characterization has utilized methods of 'reverse pharmacology', using the receptor as bait to retrieve ligands from tissue extracts. These efforts have identified endogenous ligands such as apelin, ghrelin, melanin-concentrating hormone, neuromedin U, the orexins, urotensin-II, and UDP-glucose (recently reviewed in Lee et al., 2001a; Civelli et al., 2001; Howard et al., 2001).

Approximately 250 mammalian genes encoding family A (or rhodopsin-like) GPCRs have been cloned (Lee et al., 2001a). As yet, the total number of cloned GPCRs reported in the literature including the secretin and metabotropic glutamate-like families of GPCRs falls short of the projected 616 GPCR-encoding sequences observed from the complete human genome sequence (Venter et al., 2001). Despite the human genome sequencing efforts, much work is still required to identify and clone the open reading frames (ORF) encoding the full complement of GPCR genes. Inserted into suitable expression vectors, these DNA sequences can be used to express the receptor in assays which will assist in ligand identification. For these reasons, we are continuing in our efforts to identify, catalog, compare and map the expression of GPCRs. We have recently reported the identification of the H4 histamine receptor (Nguyen et al., 2001), the cysteinyl leukotriene 2 receptor (Heise et al., 2000) and the oGPCR-encoding genes *GPR26*, *GPR57*, *GPR58* (Lee et al., 2000), *GPR61*, *GPR62*, *GPR63* and *GPR77* (Lee et al., 2001b). We now report the cloning of ten additional oGPCR-encoding genes named *GPR78*, *GPR80*, *GPR81*, *GPR82*, *GPR93*, *GPR94*, *GPR95*, *GPR101*, *GPR102* and *GPR103* as well as a pseudogene *ψGPR79*. *GPR78* and *GPR81* most closely resemble the oGPCR genes *GPR26* and *HM74*, respectively. *GPR80*, *GPR93* and *ψGPR79* shared highest identities with members of the purinoceptor family, while *GPR82* encoded an oGPCR distantly related to the purinoceptor-like oGPCR genes *GPR17* and *GPR34*. In addition, two novel genes *GPR93* and *GPR94* share significant identities with each other and with recently identified genes encoding the UDP-glucose (Chambers et al., 2000) and platelet ADP ($P2Y_{12}$) receptors (Hollopeter et al., 2001; Zhang et al., 2001), which together comprise a clustered family of genes on chromosome 3. *GPR101* shared distant identity with amine-binding GPCR genes, *GPR102* shared identity with the *PNR/GPR57/GPR58* amine receptor-like subfamily of GPCR genes and *GPR103* shared identities with peptide-binding receptors, including the neuropeptide FF 2, neuropeptide Y2 and galanin GalR1 receptors. We have also detected mRNA transcripts in tissues for *GPR78*, *GPR81*, *GPR94*, *GPR95*, *GPR101* and *GPR103*.

2. Materials and methods

2.1. Database searching

We queried the expressed sequence tag (EST) and high-throughput genomic sequences (HTGS) databases maintained by the National Center for Biotechnology Information with the amino acid (aa) sequences of various GPCRs using the TBLAST algorithm (Altschul et al., 1997). Retrieved sequences having statistically significant scores were further examined. The conceptualized protein sequences encoded by these sequences were used to query the non-redundant (nr) database to determine whether these sequences encoded previously known GPCRs.

2.2. GPCR gene and cDNA cloning

GPR78 was originally obtained in two fragments in an HTGS sequence from human chromosome 4 (GenBank Accession number: AC007104) which encoded the start methionine to the third intracellular loop (IC3) and from the carboxyl region of TM6 to the stop codon. Based on this sequence, two DNA fragments encoding from TM1 to TM4 and from TM6 to the stop codon were amplified from human genomic DNA. PCR products were extracted with phenol and chloroform, precipitated with ethanol and electrophoresed on a low-melting point agarose gel. Products in the expected size range were ligated into the *Eco*RV site of pBluescript SK(−) (Stratagene, La Jolla, CA) or $pcDNA_3$ (Invitrogen, Carlsbad, CA) and sequenced. Both fragments were observed to be identical to the HTGS sequence, and were used as probes to screen a human genomic library as previously described (Marchese et al., 1994). Library screening retrieved two phage DNA which encoded from the start methionine to IC3 and from the carboxyl region of TM6 to the stop codon. Primers designed upon TM5 (P1: 5′-GCTTCGTGCTGCCGCTG-3′) and TM7 (P2: 5′-CGGAG-CAGAGAGTACGTG-3′) were used to PCR amplify Marathon ready human fetal cDNA (Clontech, Palo Alto, CA) to retrieve a fragment sharing 100% identity in regions of overlap with the HTGS and human genomic phage DNA sequences. To obtain the complete intronless ORF of this gene, three overlapping segments encoding *GPR78* were obtained by PCR. Fragment 1 (encoding from the start methionine to TM5) was amplified from human genomic DNA using two primers (P3: 5′-GCGCCATGGGCCCCG-GCGAGG-3′, P4: 5′-GGTGACGGTGTCCATGCGC-3′). Fragment 2 was amplified using primers P1 and P2 from human fetal cDNA (described above). Fragment 3 (encoding from the third extracellular loop and extending 3′ of the stop codon) was amplified from human genomic DNA using two primers (P5: 5′-CTGGCGGAGCTCGTGCCC-3′, P6: 5′-GGCCAGTGCCCTTTCCAC-3′). These DNA fragments were joined by two further rounds of PCR. Round one consisted of fragments 1 and 2 together undergoing 30 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min using

primers P2 and P3. A second round of PCR amplified an aliquot of the first round with fragment 3 using primers P3 and P6 at similar cycling conditions. The PCR products were subcloned into the *Eco*RV site of pcDNA$_3$ (Invitrogen) and sequenced.

GPR103 was originally obtained in two overlapping fragments from the EST database (encoding from TM4 to TM7, GenBank Accession number: AI307658) and the HTGS database (encoding from TM6 to the stop codon, GenBank Accession number: AC005961). Analysis of the EST clone obtained from the I.M.A.G.E. Consortium revealed that this fragment encoded the receptor from TM2 to TM7. Based on these sequences, two DNA fragments encoding from TM3 to TM7 and from TM7 to the stop codon were obtained by PCR from human hypothalamus cDNA (Clontech) and human genomic DNA, respectively. PCR products were purified and ligated into pBluescript as described above. Sequence analysis revealed both DNA fragments to be identical with their respective database sequences. The fragment encoding TM3 to TM7 was used as a probe to screen a human hypothalamus cDNA library (Clontech), which retrieved a phage encoding *GPR103* from the start methionine to TM3 sharing 100% identity in the overlapping region with the EST derived sequence. To obtain a complete intronless ORF of this gene, the three overlapping fragments were joined by PCR as described above. The PCR products were subcloned into the *Eco*RV site of pcDNA$_3$ and sequenced.

To obtain DNA encoding other GPCRs, human genomic DNA was amplified by PCR using the following oligonucleotides: *ψGPR79* (5′-TGGGGCAGAGGCTGATGCCATGC-3′, 5′-AGCTGGATGCTCACCCAACTTGTTC-3′), *GPR80* (5′-GATTCATATTGCCAAACTGAAC-3′, 5′-CATCCTGAACATCTTAGGATG-3′), *GPR81* (5′-CTAACGCTCAGATAAGCATCTGTG-3′, 5′-GTCACCACTCTATCTTCCTCAGTG-3′), *GPR82* (5′-AATTCTATTCTAGCTCCTGTG-3′, 5′-CTAATAAAGTCACATGAATGC-3′), *GPR93* (5′-TTTGGCACGATGTTAGCC-3′, 5′-GTTCAGAGGGCGGAATCC-3′), *GPR94* (5′-AAGCAATGAACACCACAGTGATGC-3′, 5′-ATTATCTACGGAAGTCTCATC-3′), *GPR95* (5′-AGTTGGGTCTGTAAGGGAACC-3′, 5′-TTTATTTACACTTTGTACATATCG-3′), *GPR101* (5′-CTGGCTGTTGCCATGACGTCC-3′, 5′-GCCTTAGAACTAACTTCAAGG-3′), *GPR102* (5′-CAAACAACAAACAGCAGAACC-3′, 5′-CTTAGTGCTTAAACTTATTC-3′) and *P2Y$_{12}$* (5′-AAATAACCATCCTCTCTTTTGTTC-3′, 5′-CGAGTTCTGAACACAAAGAGATTG-3′). PCR conditions were as follows: denaturation at 94°C for 30 s, annealing at 50°C for 30 s, and extension at 72°C for 1 min. PCR products were purified, ligated into pcDNA3 or pBluescript vectors and sequenced as described above.

2.3. Northern blot analyses

Human and rat mRNA were extracted from various tissues as described previously (Marchese et al., 1994). Briefly, total RNA was extracted by the method of Chomczynski and Sacchi (1987), and poly (A)$^+$ RNA was isolated using oligo(dT) cellulose spin columns (Pharmacia, Uppsala, Sweden). RNA was denatured and separated by electrophoresis on a 1% formaldehyde agarose gel, transferred onto nylon membrane and immobilized by UV irradiation. The blots were hybridized with human GPCR-encoding ^{32}P-labeled DNA fragments, washed with 2× SSPE (3 M NaCl, 0.2 M NaH_2PO_4, 0.02 M EDTA) and 0.1% SDS at 50°C for 20 min, washed again with 0.1× SSPE and 0.1% SDS at 50°C for 2 h and exposed to X-ray film at −70°C in the presence of an intensifying screen. In addition, two human Multiple Tissue Northern (MTN™) blots (Clontech) were used in *GPR78* (Human MTN Blot) and *GPR95* (Human MTN Blot IV) expression analyses according to the manufacturer's instructions.

3. Results

3.1. Cloning of GPCR-encoding genes

A search of the HTGS database with the *GPR26* sequence retrieved a novel human genomic sequence encoding a GPCR localized to chromosome 4 (GenBank Accession number: AC007104). The sequence was encoded on two fragments, from the start methionine to the middle region of IC3 and from the carboxyl region of TM6 to the stop codon. Primers were designed to PCR amplify human genomic DNA which retrieved fragments encoding TM1 to TM4 and from TM6 to the stop codon. These fragments were used to screen a human genomic library and two phage DNA fragments were retrieved. We also used TM5- and TM7-specific primers to amplify a DNA fragment from human fetal cDNA. The cDNA PCR product revealed 100% identity in regions overlapping the fragments previously obtained from genomic DNA, confirming these fragments as segments of the same gene. To obtain the full-length ORF, the three overlapping fragments were joined by PCR, and this clone was named *GPR78* (Fig. 1). *GPR78* encoded a 363 aa protein which shared extensive identities in the TM regions with *GPR26* (56%) (Table 1).

A customized search of the HTGS database retrieved a human genomic sequence (GenBank Accession number: AC021773) apparently encoding a novel GPCR. However, the sequence (*ψGPR79*) contained a frame-shift in the ORF. Primers were designed and used to PCR amplify this region of DNA sequence, and the product sequenced to verify the frame-shift in the TM5-encoding region confirming a pseudogene (*ψGPR79*). *ψGPR79* was used to search the HTGS database, which retrieved a related GPCR-encoding human genomic sequence localized to chromosome 13 (GenBank Accession number: AC026756). The PCR product containing this gene revealed an ORF of 336 aa in length, which we named *GPR80*. A search of the nr database with the projected *ψGPR79* aa sequence revealed significant identi-

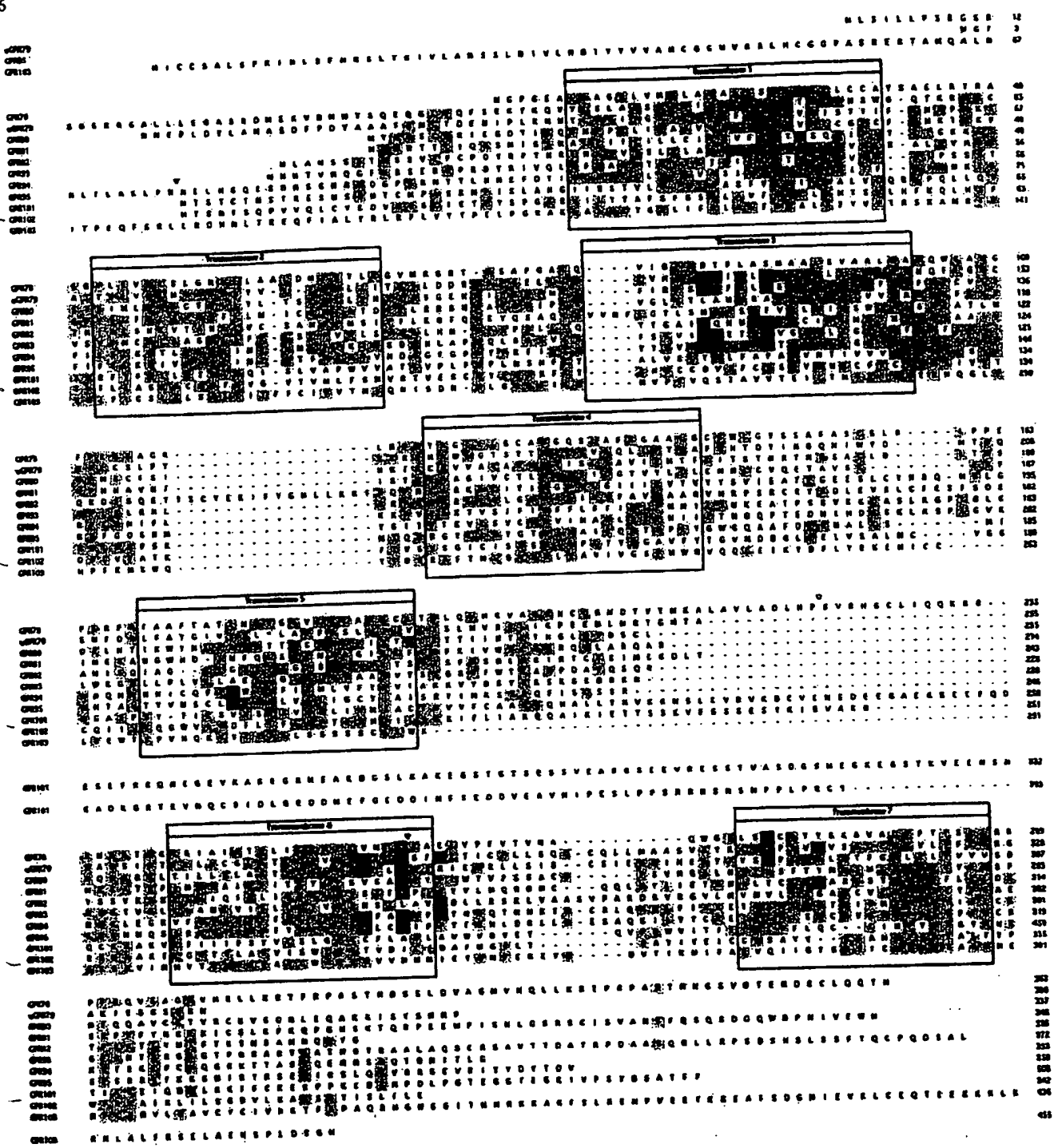


Fig. 1. oGPCR sequence alignments. Amino acid comparisons between novel GPCRs. The TM domain regions are indicated, and amino acids are numbered on the right. Black boxes with white lettering indicate conserved residues for purinergic binding (Erb et al., 1995; Jiang et al., 1997; Hoffmann et al., 1999). The presence of introns interrupting the *GPR78* and *GPR95* gene sequences is indicated by '∇' and '▼', respectively. The presence of a frame-shift interrupting the *ψGPR79* sequence is indicated by '*'. Residues shared between at least three aligned GPCR sequences are shaded.

ties in the TM regions with the purinoceptors $P2Y_2$ (51%), $P2Y_4$ (50%) and $P2Y_6$ (43%). A similar search with the *GPR80* sequence revealed greatest identities in the TM regions with the purinoceptors $P2Y_1$ (45%) and $P2Y_4$ (39%) and the cysteinyl leukotriene receptor CysLT2 (39%).

From the HTGS database we retrieved a GPCR-encoding sequence contained on a bacterial artificial chromosome (BAC) clone localized to chromosome 12q (GenBank Accession number: AC026331). The PCR amplified product (*GPR81*) encoded a 347 aa protein with identities in the TM regions with HM74 (70%), GPR31 (43%) and the purino-

Table 1
oGPCR sequence identities[a]

Receptor	% Identity	Accession number	Receptor	% Identity	Accession number
GPR78	GPR26, 49 (56) SSTR4, 21 (29) $5HT_6$, 20 (25)	AF411107	GPR95	UDP-glucose, 42 (62) $P2Y_{12}$, 39 (54) GPR94, 37 (51)	AF11114
GPR80	$P2Y_1$, 31 (45) $P2Y_4$, 29 (39) CysLT2, 31 (39)	AF411109	GPR101	RE2, <20 (31) $5HT_{1A}$, <20 (29) α_{1A}, <20 (29)	AF411115
GPR81	HM74, 50 (70) GPR31, 30 (43) $P2Y_1$, 23 (37)	AF411110	GPR102	PNR, 40 (42) GPR57, 36 (35) GPR58, 33 (35)	AF411116
GPR82	GPR34, 23 (31) GPR17, 24 (30) SSTR2, 23 (30)	AF411111	GPR103	NPFF2, 31 (38) NPY2, 29 (37) GalR1, 30 (35)	AF411117
GPR93	$P2Y_5$, 31 (40) GPR23, 28 (38) GPR17, 25 (36)	AF411112	ψGPR79	$P2Y_2$, 37 (51) $P2Y_4$, 37 (50) $P2Y_6$, 33 (43)	AF411108
GPR94	$P2Y_{12}$, 46 (57) UDP-glucose, 41 (52) GPR95, 37 (51)	AF411113			

[a] % Identity represents the shared sequence identities of each receptor with three of the closest GPCR sequences (TM identities are in parentheses). Each novel oGPCR sequence can be accessed through GenBank (http://www2.ncbi.nlm.nih.gov/genbank/query_form.html) with the listed Accession numbers.

ceptor P2Y1 (37%). The *HM74* sequence was retrieved in the same BAC clone as *GPR81*, indicating a clustering of these two genes.

GPR82 was retrieved using *GPR34* (from a search of the HTGS database) on a human BAC clone localized to chromosome 1 (GenBank Accession number: AL161458). The *GPR82* PCR product encoded a 336 aa protein sharing identities in the TM regions with the oGPCRs GPR34 (31%) and GPR17 (30%) and the somatostatin receptor SSTR2 (30%).

A search of the HTGS database using the cysteinyl leukotriene 2 receptor sequence retrieved a sequence (*GPR93*) encoding a GPCR on a human BAC clone localized to chromosome 12 (GenBank Accession number: AC006087). The PCR product obtained encoded a 372 aa protein which shared identities in the TM regions with the oGPCRs $P2Y_5$ (40%), GPR23 (38%) and GPR17 (36%).

A search of the HTGS database retrieved a cluster of three genes encoding GPCRs within the same contig localized to chromosome 3 (GenBank Accession number: AC024886). One of these genes was recently reported to encode the platelet ADP receptor, $P2Y_{12}$ (Zhang et al., 2001; Hollopeter et al., 2001). The $P2Y_{12}$ receptor gene shared homology with the other two genes, which we named *GPR94* and *GPR95*. A PCR product (*GPR94*) encoded a 333 aa protein which shared identities in the TM regions with the $P2Y_{12}$ receptor (57%), the UDP-glucose receptor (52%) and the receptor encoded by *GPR95* (51%). *GPR95* encoded a truncated GPCR, from TM1 to the stop codon, with an intron in the ORF evident by the lack of start methionine and the presence of an upstream in-frame stop codon. A search of the EST database retrieved one EST sequence from human testis mRNA encoding *GPR95* which also lacked a start methionine codon (GenBank Accession number: AA758208). This EST was acquired from the I.M.A.G.E. Consortium, sequenced, and confirmed to share 100% identity with the genomic sequence. A more recent database search retrieved an EST sequence from a human bladder cell line encoding *GPR95* with an alternative 5′ coding region (GenBank Accession number: BF028445), confirming a complete *GPR95* ORF and the presence of an intron between the start methionine and TM1-encoding regions. *GPR95* encoded a 358 aa protein which shared highest identities in the TM regions with the UDP-glucose receptor (62%), the $P2Y_{12}$ receptor (54%) and the receptor encoded by *GPR94* (51%).

We used the histamine H1 and H4 receptor sequences to retrieve two GPCR-encoding sequences, *GPR101* and *GPR102*, respectively, from the HTGS database localized to chromosome X (GenBank Accession number: AL390879) and chromosome 6 (GenBank Accession number: AL357505), respectively. A PCR amplified product (*GPR101*) encoded a 508 aa protein with identities in the TM regions with the oGPCR RE2 (31%), the serotonin $5HT_{1A}$ receptor (29%) and the α_{1A} adrenergic receptor (29%). The *GPR102* PCR product encoded a 342 aa protein sharing identities in the TM regions with the oGPCRs PNR (42%), GPR57 (35%) and GPR58 (35%).

A search of the EST and HTGS databases with the rat leukotriene LT2 receptor sequence retrieved human DNA sequences encoding a novel GPCR, GPR103. The novel receptor sequence was encoded on two overlapping fragments from a human kidney EST (encoding from TM2 to TM7) and from a human HTGS sequence (encoding from TM6 to the stop codon). Primers were designed to PCR amplify human hypothalamus cDNA and genomic DNA which retrieved two fragments encoding TM3 to TM7 and

from TM6 to the stop codon, respectively. The TM3 to TM7-encoding fragment was used to screen a human hypothalamus cDNA library, which retrieved a phage DNA fragment encoding GPR103 from the start methionine to TM3. The three DNA fragments revealed 100% identity in regions of overlap, confirming these fragments as segments of the same gene. To obtain the full-length ORF, the three overlapping fragments were joined by PCR, and this clone was named *GPR103*. *GPR103* encoded a 455 aa protein which shared identities in the TM regions with various peptide receptors, including the neuropeptide FF 2 (38%), neuropeptide Y2 (37%) and galanin GalR1 (35%) receptors.

3.2. Expression analyses

GPR78 mRNA transcripts were detected in human pituitary (1.1 kb) and placenta (two signals of 4.2 and 1.1 kb in size) (Fig. 2A,B). However, no *GPR78* transcripts were observed in human brain or specific CNS regions such as the frontal cortex, putamen, thalamus, hypothalamus, amygdala, hippocampus, pons, medulla and midbrain. In addition, human *GPR78* transcripts were also absent from skeletal muscle, lung, heart, liver, pancreas and kidney.

To determine the expression distribution of these novel oGPCRs, we performed Northern blots in various human and rat tissues. *ψGPR79* mRNA transcripts were not detected in human brain tissues, including frontal cortex, basal forebrain, pituitary, caudate nucleus, nucleus accumbens or hippocampus. In addition, Northern analyses did not reveal *ψGPR79* in rat brain, fetus, liver, spleen or adrenal gland tissue. Similarly, *GPR80* mRNA transcripts were not detected in human brain tissues including the frontal cortex, caudate putamen, thalamus, hypothalamus, hippocampus or pons tissue.

An mRNA transcript was detected for *GPR81* in human pituitary tissue, with an absence of signal in frontal, temporal and occipital lobes of the cortex, basal forebrain, caudate nucleus, nucleus accumbens, and hippocampus (Fig. 2C). Analyses of *GPR82* mRNA expression revealed no transcripts in human tissues, including regions of the CNS such as the frontal cortex, caudate putamen, thalamus, hypothalamus, hippocampus, pons and liver tissue. A search of the GenBank database retrieved an EST from cells derived from human colonic tissue encoding *GPR82* (GenBank Accession number: BF335802). The *GPR93* probe failed to detect mRNA transcripts in human frontal cortex, basal forebrain, caudate putamen, thalamus, or hippocampus. A search of the GenBank database revealed an EST encoding *GPR93* from human primary tonsil B-cells (GenBank Accession numbers: BF975186, BF663176 and BF129117) and an EST encoding a mouse *GPR93* orthologue expressed in the small intestine (GenBank Accession numbers: AV064817 and AV064680).

GPR94 mRNA transcripts of 3.2 kb were detected in human CNS tissues including the frontal cortex, caudate

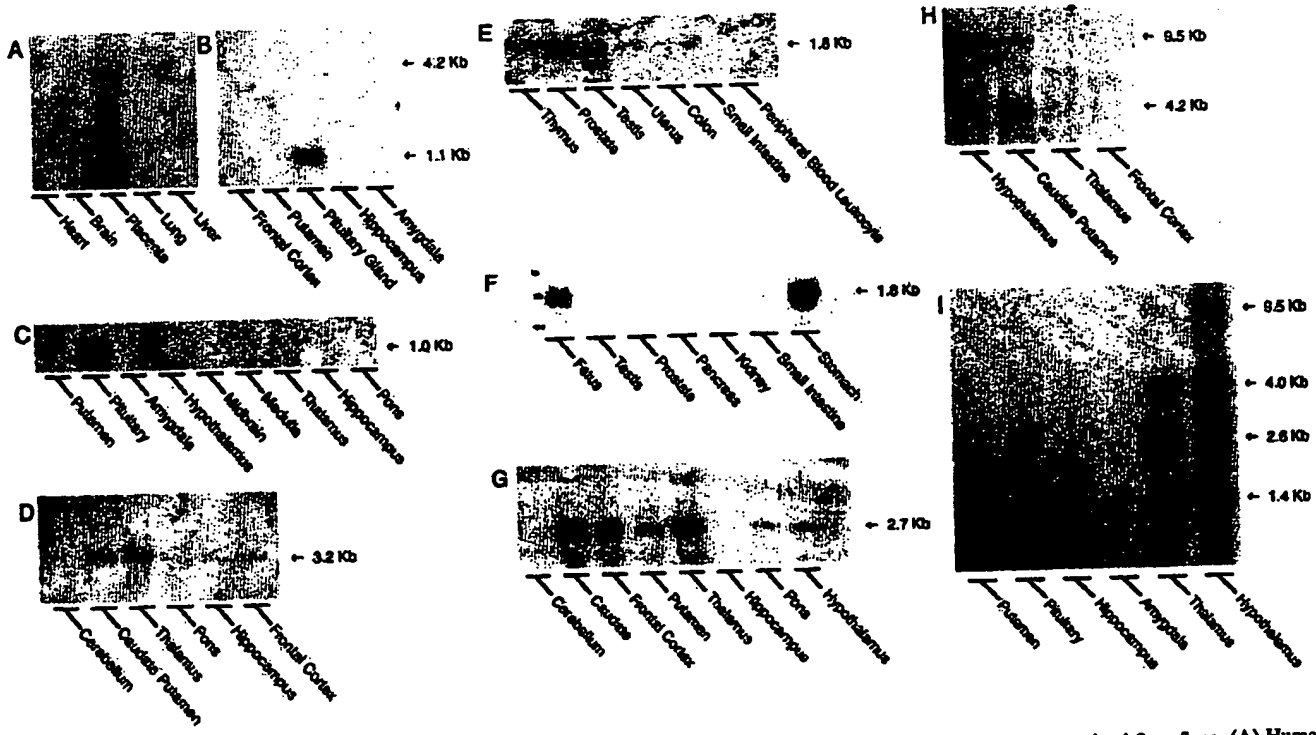


Fig. 2. oGPCR tissue distribution analyses. Northern blot of poly(A)$^+$ RNA (10 μg/lane), except for (A) and (E) which contained 2 μg/lane. (A) Human MTN™ and (B) human tissue distribution of *GPR78*. (C) *GPR81* and (D) *GPR94* human CNS tissue distribution. (E) Human MTN™ and (F) rat tissue distribution of *GPR95*, hybridized with a radiolabeled fragment encoding human *GPR95*. (G) *P2Y$_{12}$* and (H) *GPR101* human CNS tissue distribution. (I) Human tissue distribution of *GPR103*.

putamen and thalamus (Fig. 2D). *GPR94* was not detected in the hippocampus, pons or cerebellum and $P2Y_{12}$ was not detected in the hippocampus, cerebellum or in peripheral liver tissue. *GPR95* transcripts were observed in both human and rat peripheral tissue. A 1.8 kb signal was detected in human prostate, rat stomach and rat fetal tissues (Fig. 2E,F). However, *GPR95* expression was not observed in human thalamus, hypothalamus, hippocampus, pons or cerebellum or in rat whole brain tissue. In addition, we performed Northern analyses of $P2Y_{12}$ transcripts in human CNS tissues, which revealed faint signals of 2.7 kb in the frontal cortex, caudate putamen, thalamus, hypothalamus and pons (Fig. 2G). *GPR101* mRNA transcripts of 9.5 and 4.2 kb were detected in the caudate putamen and hypothalamus, with no expression detected in the frontal cortex, thalamus, hippocampus and pons (Fig. 2H). *GPR103* transcripts of 4.0, 2.6 and 1.4 kb were detected in the thalamus and hypothalamus, with a further 9.5 kb signal in the hypothalamus and a 1.4 kb signal in the pituitary (Fig. 2I). In addition, *GPR103* transcripts were also observed in the frontal and occipital cortices, basal forebrain, midbrain and pons (data not shown).

4. Discussion

Currently, ~350 human GPCRs have been cloned, as listed on the GPCRDB (G protein-coupled receptor database, http://www.gpcr.org/7tm/), with ~250 representing family A (or rhodopsin-like) GPCRs (Lee et al., 2001a). These receptors total approximately half the predicted 616 GPCR-encoding sequences contained in the human genome (Venter et al., 2001), although the veracity of this total number remains to be confirmed.

The identification of genes encoding the novel GPCRs predicts the existence of novel signaling systems leading to the discovery of novel ligands, as demonstrated by recent reports describing the discovery of apelin (Tatemoto et al., 1998), prolactin-releasing peptide (Hinuma et al., 1998), orexin (Sakurai et al., 1998), melanin-concentrating hormone (Bachner et al., 1999; Chambers et al., 1999; Lembo et al., 1999; Saito et al., 1999; Shimomura et al., 1999) and urotensin II (Ames et al., 1999; Liu et al., 1999; Mori et al., 1999; Nothacker et al., 1999) receptor-ligand systems. As these GPCR genes (many of which were cloned in our laboratory) were used in methods that led to the discovery and identification of these ligands (Lee et al., 2001a), we are continuing to isolate and characterize these novel genes.

We now report the discovery of ten novel oGPCRs and a pseudogene. *GPR78* is a paralogue of *GPR26*, apparent from shared identities (56% in the TM regions), a lack of asparagine-linked extracellular glycosylation sites, a short amino terminus, and similar gene structure (Fig. 1). *GPR78* and *GPR26* encoded receptors with shared conserved cationic arginine and lysine residues in TM6 and TM7, respectively, two residues recognized to play a role in purinergic binding and found only in P2Y receptors (Erb et al., 1995; Jiang et al., 1997) (Fig. 1). However, we reported calcium mobilization assays of human astrocytoma 1321N1 cells and *Xenopus laevis* oocytes transfected with *GPR26* were not responsive to nucleoside di- and tri-phosphates (Lee et al., 2000). We reported high levels of *GPR26* expression in many brain regions (Lee et al., 2000), while *GPR78* was detected only in the pituitary and placenta. However, the overall structural homology suggested that *GPR78* and *GPR26* may encode receptors that share a common endogenous ligand.

GPR80, *GPR81*, *GPR82*, *GPR93* and *ψGPR79* all shared identities to P2Y GPCRs or P2Y-like oGPCRs. Previously, at least three different nucleotide receptor phenotypes have been observed in mammalian tissue, including GPCRs activated by adenine nucleotides (e.g. $P2Y_1$ and $P2Y_{11}$), uridine nucleotides (e.g. $P2Y_6$) and by both adenine and uridine (e.g. $P2Y_2$ and $P2Y_4$) (King et al., 1998). *ψGPR79* shared closest identity with $P2Y_2$ and $P2Y_4$, even though it does not encode a functional GPCR. *GPR80* was observed to share highest identity with $P2Y_1$, while *GPR81*, *GPR82* and *GPR93* shared identities with the P2Y-like oGPCR genes *HM74*, *GPR34* and $P2Y_5$, respectively. Some of these oGPCR genes encode aa residues conserved amongst the P2Y receptors and shown to be involved in purinergic ligand binding (Erb et al., 1995; Hoffmann et al., 1999; Jiang et al., 1997) (Fig. 1). While expression was not observed for *GPR80*, *GPR82* and *GPR93* in various CNS regions, *GPR81* was observed to have an mRNA transcript in the pituitary, suggesting a role in neuroendocrine regulation.

The identities of *GPR94* and *GPR95* with the genes encoding the UDP-glucose and $P2Y_{12}$ receptors (>50% in the TM regions) indicate a novel subfamily of purinergic-like receptors. Previously, the UDP-glucose receptor was reported to have a distant sequence homology with the P2Y receptors, with an observed widespread tissue distribution in human brain and such peripheral tissues as placenta, adipose tissue, spleen, intestine, stomach, skeletal muscle, lung and heart (Chambers et al., 2000). The identification of the platelet ADP ($P2Y_{12}$) receptor resulted from cDNA isolated from rat platelets and human hypothalamus (Zhang et al., 2001; Hollopeter et al., 2001). An alignment of these receptors with the GPCRs encoded by *GPR94* and *GPR95* revealed several residues conserved in P2Y purinoceptors (see above) also conserved throughout this novel purinoceptor-like subfamily (Fig. 1). Northern analysis of *GPR94* mRNA revealed expression in various regions of the brain, suggesting a neuromodulatory role. In contrast, *GPR95* mRNA was detected in peripheral tissue (i.e. human prostate and rat stomach).

The receptor encoded by *GPR101* appeared to be a distant relative of the biogenic amine superfamily of GPCRs, with TM identities of ~30% with the adrenergic and serotonin receptors, as well as the muscarinic and dopamine receptors (data not shown). *GPR101* mRNA transcripts were observed in brain tissue, suggesting the presence of an endogenous

amine neurotransmitter ligand, perhaps novel in identity. The receptor encoded by *GPR102* shared significant TM identities with an amine binding receptor-like GPCR family (including PNR, GPR57 and GPR58) suggesting they may also share a common endogenous ligand. The receptor encoded by *GPR103* shared highest identities with several neuropeptide receptors. The significant levels of *GPR103* expression in the brain, particularly in the thalamus and hypothalamus, suggest an endogenous peptide ligand, perhaps involved in physiological functions such as pain modulation and neuroendocrine regulation.

Interestingly, some of these novel oGPCR genes appear to be clustered on various human chromosomes. The *GPR81* gene was localized to chromosome 12q, proximal to the closely related oGPCR gene, *HM74*. The *GPR82* gene was retrieved together with the *GPR34* gene in a human BAC clone localized on chromosome 1. Another search of the HTGS database revealed *GPR94*, *GPR95* and $P2Y_{12}$ localized within the same BAC clone on chromosome 3. This cluster of genes also includes the UDP-glucose receptor gene, which together with the more distant $P2Y_1$ receptor gene further localizes this cluster to chromosome 3q24-25 (interval D3S1279-1280) (Hollopeter et al., 2001). We have previously reported clusters of homologous GPCRs genes, including the *GPR40* through *GPR43* gene cluster (Sawzdargo et al., 1997) as well as the 5-HT_4-like pseudogene, *ψGPR57*, *GPR58* and *PNR* gene cluster (Lee et al., 2000). Given the significant sequence similarity of *GPR102* with *PNR*, *GPR57* and *GPR58* and its localization on chromosome 6, *GPR102* may be another paralogue member of this gene cluster.

In conclusion, we have identified ten novel GPCR genes and a pseudogene. Transcripts for *GPR78*, *GPR81*, *GPR94*, *GPR95*, *GPR101*, *GPR103* and $P2Y_{12}$ were detected in various CNS and peripheral tissues. Given the high levels of identity observed within paralogous oGPCR gene clusters, future efforts will likely discover common endogenous ligands for each of these novel GPCR subfamilies. The increasing number of oGPCRs that continue to be isolated with unique distribution profiles in brain and periphery is indicative that the search for novel transmitter ligands should be intensified. These efforts have the tremendous potential to uncover novel physiological roles for these as yet unknown receptor-transmitter signaling systems.

Acknowledgements

This study was funded by a Canadian Institutes of Health Research (CIHR) grant to S.R.G. and B.F.O. and by Merck Frosst Research Laboratories, Inc.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Ames, R.S., Sarau, H.M., Chambers, J.K., Willette, R.N., Aiyar, N.V., Romanic, A.M., Louden, C.S., Foley, J.J., Sauermelch, C.F., Coatney, R.W., Ao, Z., Disa, J., Holmes, S.D., Stadel, J.M., Martin, J.D., Liu, W.S., Glover, G.I., Wilson, S., McNulty, D.E., Ellis, C.E., et al., 1999. Human urotensin-II is a potent vasoconstrictor and agonist for the orphan receptor GPR14. Nature 401, 282–286.

Bachner, D., Kreienkamp, H., Weise, C., Buck, F., Richter, D., 1999. Identification of melanin concentrating hormone (MCH) as the natural ligand for the orphan somatostatin-like receptor 1 (SLC-1). FEBS Lett. 457, 522–524.

Chambers, J., Ames, R.S., Bergsma, D., Muir, A., Fitzgerald, L.R., Hervieu, G., Dytko, G.M., Foley, J.J., Martin, J., Liu, W.S., Park, J., Ellis, C., Ganguly, S., Konchar, S., Cluderay, J., Leslie, R., Wilson, S., Sarau, H.M., 1999. Melanin-concentrating hormone is the cognate ligand for the orphan G-protein-coupled receptor SLC-1. Nature 400, 261–265.

Chambers, J.K., Macdonald, L.E., Sarau, H.M., Ames, R.S., Freeman, K., Foley, J.J., Zhu, Y., McLaughlin, M.M., Murdock, P., McMillan, L., Trill, J., Swift, A., Aiyar, N., Taylor, P., Vawter, L., Naheed, S., Szekeres, P., Hervieu, G., Scott, C., Watson, J.M., et al., 2000. A G protein-coupled receptor for UDP-glucose. J. Biol. Chem. 275, 10767–10771.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156–159.

Civelli, O., Nothacker, H., Saito, Y., Wang, Z., Lin, S.H., Reinscheid, R.K., 2001. Novel neurotransmitters as natural ligands of orphan G-protein-coupled receptors. Trends Neurosci. 24, 230–237.

Erb, L., Garrad, R., Wang, Y., Quinn, T., Turner, J.T., Weisman, G.A., 1995. Site-directed mutagenesis of P2U purinoceptors. Positively charged amino acids in transmembrane helices 6 and 7 affect agonist potency and specificity. J. Biol. Chem. 270, 4185–4188.

Heise, C.E., O'Dowd, B.F., Figueroa, D.J., Sawyer, N., Nguyen, T., Im, D.S., Stocco, R., Bellefeuille, J.N., Abramovitz, M., Cheng, R., Williams, J., Zeng, Z., Liu, Q., Ma, L., Clements, M.K., Coulombe, N., Liu, Y., Austin, C.P., George, S.R., O'Neill, G.P., et al., 2000. Characterization of the human cysteinyl leukotriene 2 receptor. J. Biol. Chem. 275, 30531–30536.

Hinuma, S., Habata, Y., Fujii, R., Kawamata, Y., Hosoya, M., Fukusumi, S., Kitada, C., Masuo, Y., Asano, T., Matsumoto, H., Sekiguchi, M., Kurokawa, T., Nishimura, O., Onda, H., Fujino, M., 1998. A prolactin-releasing peptide in the brain. Nature 393, 272–276.

Hoffmann, C., Moro, S., Nicholas, R.A., Harden, T.K., Jacobson, K.A., 1999. The role of amino acids in extracellular loops of the human P2Y1 receptor in surface expression and activation processes. J. Biol. Chem. 274, 14639–14647.

Hollopeter, G., Jantzen, H.-M., Vincent, D., Li, G., England, L., Ramakrishnan, V., Yang, R.-B., Nurden, P., Nurden, A., Julius, D.J., Conley, P.B., 2001. Identification of the platelet ADP receptor targeted by antithrombotic drugs. Nature 409, 202–207.

Howard, A.D., McAllister, G., Feighner, S.D., Liu, Q., Nargund, R.P., Van der Ploeg, L.H., Patchett, A.A., 2001. Orphan G-protein-coupled receptors and natural ligand discovery. Trends Pharmacol. Sci. 22, 132–140.

Jiang, Q., Guo, D., Lee, B.X., Van Rhee, A.M., Kim, Y.C., Nicholas, R.A., Schachter, J.B., Harden, T.K., Jacobson, K.A., 1997. A mutational analysis of residues essential for ligand recognition at the human P2Y1 receptor. Mol. Pharmacol. 52, 499–507.

King, B.F., Townsend-Nicholson, A., Burnstock, G., 1998. Metabotropic receptors for ATP and UTP: exploring the correspondence between native and recombinant nucleotide receptors. Trends Pharmacol. Sci. 19, 506–514.

Lee, D.K., Lynch, K.R., Nguyen, T., Im, D., Cheng, R., Saldivia, V.R., Liu, Y., Liu, I.S., Heng, H.H., Seeman, P., George, S.R., O'Dowd, B.F., Marchese, A., 2000. Cloning and characterization of additional

members of the G protein-coupled receptor family. Biochim. Biophys. Acta 1490, 311–323.

Lee, D.K., George, S.R., Evans, J.F., Lynch, K.R., O'Dowd, B.F., 2001a. Orphan G protein-coupled receptors in the CNS. Curr. Opin. Pharmacol. 1, 31–39.

Lee, D.K., George, S.R., Cheng, R., Nguyen, T., Liu, Y., Brown, M., Lynch, K.R., O'Dowd, B.F., 2001b. Identification of four novel human G protein-coupled receptors expressed in the brain. Brain Res. Mol. Brain Res. 86, 13–22.

Lembo, P.M., Grazzini, E., Cao, J., Hubatsch, D.A., Pelletier, M., Hoffert, C., St-Onge, S., Pou, C., Labrecque, J., Groblewski, T., O'Donnell, D., Payza, K., Ahmad, S., Walker, P., 1999. The receptor for the orexigenic peptide melanin-concentrating hormone is a G-protein-coupled receptor. Nat. Cell Biol. 1, 267–271.

Liu, Q., Pong, S.S., Zeng, Z., Zhang, Q., Howard, A.D., Williams Jr., D.L., Davidoff, M., Wang, R., Austin, C.P., McDonald, T.P., Bai, C., George, S.R., Evans, J.F., Caskey, C.T., 1999. Identification of urotensin II as the endogenous ligand for the orphan G-protein-coupled receptor GPR14. Biochem. Biophys. Res. Commun. 266, 174–178.

Marchese, A., Docherty, J.M., Nguyen, T., Heiber, M., Cheng, R., Heng, H.H., Tsui, L.C., Shi, X., George, S.R., O'Dowd, B.F., 1994. Cloning of human genes encoding novel G protein-coupled receptors. Genomics 23, 609–618.

Marchese, A., George, S.R., O'Dowd, B.F., 1998. Cloning of G protein-coupled receptor genes: the use of homology screening and the polymerase chain reaction. In: Lynch, K.R. (Ed.), Identification and Expression of G Protein-Coupled Receptors. Wiley-Liss, New York, pp. 1–26.

Marchese, A., George, S.R., Kolakowski Jr., L.F., Lynch, K.R., O'Dowd, B.F., 1999. Novel GPCRs and their endogenous ligands: expanding the boundaries of physiology and pharmacology. Trends Pharmacol. Sci. 20, 370–375.

Mori, M., Sugo, T., Abe, M., Shimomura, Y., Kurihara, M., Kitada, C., Kikuchi, K., Shintani, Y., Kurokawa, T., Onda, H., Nishimura, O., Fujino, M., 1999. Urotensin II is the endogenous ligand of a G-protein-coupled orphan receptor, SENR (GPR14). Biochem. Biophys. Res. Commun. 265, 123–129.

Nguyen, T., Shapiro, D.A., George, S.R., Setola, V., Lee, D.K., Cheng, R., Rauser, L., Lee, S.P., Lynch, K.R., Roth, B.L., O'Dowd, B.F., 2001. Discovery of a novel member of the histamine receptor family. Mol. Pharmacol. 59, 427–433.

Nothacker, H.P., Wang, Z., McNeill, A.M., Saito, Y., Merten, S., O'Dowd, B.F., Duckles, S.P., Civelli, O., 1999. Identification of the natural ligand of an orphan G-protein-coupled receptor involved in the regulation of vasoconstriction. Nat. Cell Biol. 1, 383–385.

Saito, Y., Nothacker, H.P., Wang, Z., Lin, S.H.S., Leslie, F., Civelli, O., 1999. Molecular characterization of the melanin-concentrating-hormone receptor. Nature 400, 265–269.

Sakurai, T., et al., 1998. Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. Cell 92, 573–585.

Sawzdargo, M., George, S.R., Nguyen, T., Xu, S., Kolakowski, L.F., O'Dowd, B.F., 1997. A cluster of four novel human G protein-coupled receptor genes occurring in close proximity to CD22 gene on chromosome 19q13.1. Biochem. Biophys. Res. Commun. 239, 543–547.

Shimomura, Y., Mori, M., Sugo, T., Ishibashi, Y., Abe, M., Kurokawa, T., Onda, H., Nishimura, O., Sumino, Y., Fujino, M., 1999. Isolation and identification of melanin-concentrating hormone as the endogenous ligand of the SLC-1 receptor. Biochem. Biophys. Res. Commun. 261, 622–626.

Tatemoto, K., Hosoya, M., Habata, Y., Fujii, R., Kakegawa, T., Zou, M.X., Kawamata, Y., Fukusumi, S., Hinuma, S., Kitada, C., Kurokawa, T., Onda, H., Fujino, M., 1998. Isolation and characterization of a novel endogenous peptide ligand for the human APJ receptor. Biochem. Biophys. Res. Commun. 251, 471 476.

Venter, J.C., et al., 2001. The sequence of the human genome. Science 291, 1304–1351.

Zhang, F.L., Luo, L., Gustafson, E., Lachowicz, J., Smith, M., Qiao, X., Liu, Y.H., Chen, G., Pramanik, B., Laz, T.M., Palmer, K., Bayne, M., Monsma, F., 2001. ADP is the cognate ligand for the orphan G-protein coupled receptor SP1999. J. Biol. Chem. 276, 8608–8615.

GPCR Collection

Introduction

Product Description

Product Application

Search and Order Human Genes

Terms and Conditions

Disclaimer

Product Manual

Vector Sequences

Legal Notices

Successful Search

Accn	Clone	Description	Cost
NM_054021	MG1027_F10 [1]	Homo sapiens G protein-coupled receptor 101 (GPR101), mRNA	$80(

This price is valid only in the USA and Canada

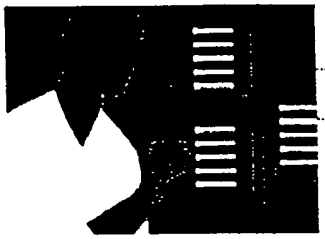
Add to Shopping Cart

[1] The 5' and 3' sequences of our molecular clones have been matche(to an accession number as a point of reference. Note that the *complet* sequence of our molecular clones may differ from the sequence published for this corresponding accession number, e.g., by representing an alternative RNA splicing form.

>MG1027_F10 OriGene 5' read for NM_054021
GCTTCTGGCTGTTGCCATGACGTCCACCTGCACCAACAGCACGCGCGAGAGTAACAGCA(
CCACACGTGCATGCCCCTCTCCAAAATGCCCATCAGCCTGGCCCACGGCATCATCCGCT(
AACCGTGCTGGTTATCTTCCTCGCCGCCTCTTTCGTCGGCAACATAGTGCTGGCGCTAG
GTTGCAGCGCAAGCCGCAGCTGCTGCAGGTGACCAACCGTTTTATCTTTAACCTCCTCG
CACCGACCTGCTGCAGATTTCGCTCGTGGCCCCCTGGGTGGTGGCCACCTCTGTGCCTC
CTTCTGGCCCCTCAACAGCCACTTCTGCACGGCCCTGGTTAGCCTCACCCACCTGTTCG(
CTTCGCCAGCGTCAACACCATTGTCTTGGTGTCAGTGGATCGCTACTTGTCCATCATCC
CCCTCTCTCCTACCCGTCCAAGATGACCCAGCGCCGCGGTTACCTGCTCCTCTATGGCA(
CTGGATTGTGGCCATCCTGCAGAGCACTCCTCCACTCTACGGCTGGGGCCAGGCTGCCT
TGATGAGCGCAATGCTCTCTGCTCCATGATCTGGGGGGCCAGCCCCAGCTACACTATTC
CAGCGTGGTGTCCTTCATCGTCATTCCACTGATTGTCATGATTGCCTGCTACTCCGTGG
GTTCTGTGCAGCCCGGAGCAGCATGCTCTGCTGTA

View the NCBI page for this accession number (in new window)
List InterPro domains that OriGene associates with this accession number

Search for another accession number



Invitrogen Clones

- Invitrogen Clones
- Search Clones
  - CloneRange™
  - Ultimate™ ORF Browser
  - Important Gene Families
- Clone Collections
  - Ultimate™ ORF Clones
  - Genestorm® Clones
  - Full-Length Clones
  - EST Clones
  - BAC/PAC Clones

Online Catalog - Invitrogen Clones

Ultimate™ ORF Browser:

Advanced Search for Ultimate™ ORF Clones

Search By ID or Keyword | Search By Sequence | Browse By Gene Ontology

33 total records for G-Protein Coupled Receptors

Buy	Clone ID	Species	Definition	Gene Symbol
☐	IOH3294	Human	complement component 5 receptor 1 (C5a ligand); complement component-5 receptor-2 (C5a ligand)	C5R1
☐	IOH12614	Human	purinergic receptor P2Y, G-protein coupled, 11	P2RY11
☐	IOH22483	Human	clone MGC:33224 IMAGE:5267661, mRNA, complete cds.	RDC1
☐	IOH14039	Human	Similar to putative nuclear protein ORF1-FL49	ORF1-FL49
☐	IOH11484	Human	glycoprotein Ib (platelet), alpha polypeptide	GP1BA
☐	IOH1987	Human	tachykinin receptor 1 isoform short; NK-1 receptor; Tachykinin receptor 1 (substance P receptor; neurokinin-1 receptor); tachykinin 1 receptor (substance P receptor, neurokinin 1 receptor); neurokinin 1 receptor	TACR1
☐	IOH13056	Human	similar to POSSIBLE GUSTATORY RECEPTOR CLONE PTE01	LOC11513:
☐	IOH9916	Human	coagulation factor II (thrombin) receptor-like 1	F2RL1
☐	IOH9624	Human	vasoactive intestinal peptide receptor 2	VIPR2
☐	IOH10679	Human	endothelin receptor type A	EDNRA
☐	IOH22637	Human	Similar to parathyroid hormone receptor 1, clone MGC:34562 IMAGE:5180885, mRNA, complete cds.	PTHR1
☐	IOH13583	Human	Duffy blood group	FY
☐	IOH4585	Human	cholecystokinin B receptor	CCKBR
☐	IOH11033	Human	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 4; G protein-coupled receptor; LPA receptor EDG4; Lysophosphatidic acid receptor EDG4	EDG4
☐	IOH10866	Human	CD97 antigen isoform 2 precursor; leukocyte antigen CD97; seven-span transmembrane protein	CD97
☐	IOH22632	Human	formyl peptide receptor-like 1; lipoxin A4 receptor (formyl peptide receptor related)	FPRL1
☐	IOH22669	Human	adrenomedullin receptor	ADMR
☐	IOH13239	Human	super conserved receptor expressed in brain 3	SREB3

FabGennix Inc.
INTERNATIONAL

Customer Service: 1800 786 1236
Technical Support: 318 219 1123
Fax: 318 798 1849
Info@fabgennix.com
www.fabgennix.com

New Item

Novel Orphan retinal G-protein coupled Receptor (GPCR-75) selective antibodies

Anti-GPCR-75Antibodies (GPCR75-100P, GPCR75-101AP and GPCR75-112AP)

Recently a novel human G-protein coupled receptor gene has been characterized and mapped to chromosome 2p16. This gene codes for a 540 amino acid protein in retinal pigment epithelium (RPE) and cells surrounding retinal arterioles. In contrast, the Northern blot data obtained from mouse sections suggest the expression of transcripts in photoreceptor inner segments and I outer plexiform layer. The transcripts of the GPCR-75 gene (7kb) are also found in abundance in brain sections. So far, no mutations in GPCR-75 protein were identified in patients suffering from Doyne's honeycomb retinal dystrophy (DHRD), an inherited retinal degeneration disease that maps to chromosome 2p16 (1).

The GPCR-75 protein is approximately 78 kDa (540 amino acids) protein that is primarily expressed in human retinal pigment epithelium (RPEs). The GPCR-75 sequence analyses suggest the presence of 7 trans-membrane domains, a characteristic feature of GPCR. The protein has a large 3 intra cellular loop which has putative N-glycosylation sites near the extra cellular N-terminal end of the proteins. The protein might be the site for interaction of G-proteins. The short carboxy terminal is intracellular and has putative post-translational modification lipid modification sites.

The Anti-GPCR-75-selective antibodies were generated against conserved sequences near N- and C-termini of the protein that are unique to GPCR-75 protein. The polyclonal antibody strongly labels a 78 kDa protein in RPE cell extracts. Anti-GPCR-75-selective antibody is also available in affinity-purified form for confocal, Western blotting and immunocytochemical analyses. *FabGennix Int. Inc.* will also conjugate antibodies with fluorescent probes upon request at extra charge. *FabGennix Int. Inc.* will also provides antibodies against proteins that are involved in retinal degenerative diseases such as various Anti-PDE antibodies, Anti-MERTK, Anti-Phospho-MERTK, EGF-containing fibulin like intracellular protein (EFEMP1), Anti-Myocilin (TIGR), Anti-Bestrophin, Anti-ELVOL4 and a Usher syndrome specific Anti-USH2a antibodies etc. FabGennix Int. Inc employs cyclic peptide methodology for generating antibodies, which results in higher titer and specificity (2). FabGennix Int. Inc., will also provide Western blot positive controls for most of these antibodies in ready-to-use buffer for easy identification of respective proteins. Limited quantities of antigens are also available. Please enquire for their availability before ordering.

Catalog #	Host Species	Nature	Cross reactivity	Quantity	volume	Price
GPCR75-100P	Rabbit	Polyclonal antisera	R, M, H	100 ml	100 ul	$ 195.00
GPCR75-101AP	Rabbit	Affinity purified IgG	R, M, H	100 ug	150 ul	$ 225.00
GPCR75-112AP	Rabbit	Affinity purified IgG	R, M, H	100 ug	150 ul	$ 225.00
PC-GPCR75	N/A	WB positive control	Rat	For 5 App	60 ul	$ 75.00
P-GPCR75	N/A	Antigenic peptides	n/a	250 ug	inquire	$ 65.00

R = rat; M = mouse; H = human; C = chicken; monk = monkey ; * not all variants are labeled equally

Immunogen: Synthetic cyclic peptide (GPCR75-101AP = PNATSLHVPHSQEGNSTS-amide; GPCR75-112AP = STSLQEGLQDLIHTATLVTC-amide).

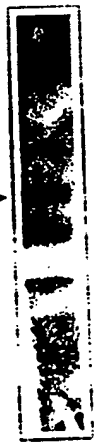
Concentration: GPCR75-101AP; GPCR-112AP IgG concentration 0.75-1.25 mg/ml in 50% antibody stabilization buffer.

Applications: Antibody GPCR75-100/GPCR75-101AP are ideal for WB, IMM and IHC assays. The dilutions for this antibody is for reference only, investigators are expected to determine the optimal conditions for specific assay in his/her laboratory. Dilutions: WB > 1:500; Immunoprecipitation & i.p pull-down assays> 1:250

Reactivity: This antibody detects a single 78 kDa Orphan GPCR75 protein in human RPE cell extracts.

Protocols: Standard protocol for various applications (WB; IMM and IHC) of this antibody is provided with the product specification sheet, however, FabGennix Int. Inc. strongly recommends investigators to optimize conditions for use of this antibody in their laboratories.

Form/Storage: The antiserum is supplied in antibody stabilization buffer with 0.02% sodium azide or thimerosal/merthiolate as preservative. The affinity-purified antibodies are purified on antigen-spepharose affinity column and supplied as 1-1.25 mg/ml IgG in antibody stabilization buffer containing preservatives with low viscosity and cryogenic properties. For long-term storage of antibodies, store at -20°C. Now these antibodies can be stored at -20°C and used immediately with out thawing. FabGennix Inc. does not recommend storage of very dilute antibody solutions unless they are prepared in specially formulated mufti use antibody dilution buffer (Cat # DiluOBuffer). Working solutions of antibodies in DiluOBuffer should be filtered through 0.45µ filter after every use for long-term storage.



78 kDa Orphan Receptor-75 in human RPE cells. Antibody GPCR-100P (1:400)

References:

1. Tarttelin E. E., Krischner L. S., Bellingham J., Baffi. J. Taymanas S. E., Gregor E. K., Csaky K., Stratakis C. A., Gregory-Evans C. Y. Biochem. Biophys. Res. Commun. 260, 174-180, 1999.
2. Farooqui, S. M., Brock. W. J., A. Hamdi., Prasad. C. (1991) J. Neurochem. 57, 1363-1369.

* For users who may require large amounts of GPCR75-100P or GPCR75-101AP, please enquire about bulk material discounts.

This Product is for Research Use Only and is NOT intended for use in humans or clinical diagnosis.

061901-0020SF1001Z-rev10.00

FabGennix Inc.
INTERNATIONL
2940 Youree Drive, Suite E, Shreveport, LA 71104

生工有限公司

Rat Taste Receptor 2 (TR2) Antibodies

Rat Taste Receptor 2 (TR2) Antibodies

Cat. # TR21-P, Rat TR2 Control Peptide # 1, SIZE: 100 ug/100 ul
FORM: Œ Soln Œ Lyophilized Lot # 3113P

Cat. # TR21-S, Rabbit Anti-rat TR2 antiserum # 1, SIZE: 100 ul neat antiserum
FORM: Œ Soln Œ Lyophilized. Lot # 38889S

Cat. # TR21-A, Rabbit Anti-rat TR2 Ab # 1 (affinity pure) SIZE: 100 ug
FORM: Œ Soln Œ Lyophilized. Lot # 38889A

Higher vertebrates are believed to possess at least five basic tastes: Sweet, bitter, sour, salty, and unami (the taste of monosodium glutamate). Taste receptor cells that may selectively reside in various parts of the tongue and respond to different tastants and perceive these taste modalities. Circumvallate papillae, found at the very back of the tongue, are particularly sensitive to biter substances. Foliate papillae, found at the posterior lateral edge of the tongue, are sensitive to sour and bitter. Fungiform papillae at the front of the tongue specialize in sweet taste.

Recently, two novel taste receptors, TR1 and TR2, have been cloned with distinct topographical distribution in taste receptor cells and taste buds. TRs are members of a new group of 7 TM domain containing GPCR distantly related to other chemosensory receptors (Ca+-sensing receptor (CaSR, a family of putative hormone receptor (V2R), and metabotropic glutamate receptors). TR1 is expressed in all fungiform taste buds, whereas TR2 localized to the circumvallate taste buds. Both receptors do not co-localize with gustducin.

Source of Antigen and Antibodies

TR1 (rat 840 aa) and TR2 (rat 843 aa) share ~40% homology with each other, and ~30% with CaSR, and 22-30% with V2R pheromone receptors and mGLURs. Rat TR are 7 TM domain containing protein with an extra long N-terminal, extracellular domain (1). A 19 AA Peptide (designated TR21-P; control peptide) sequence near the C-terminus of rat TR2(1) was selected for antibody production. The peptide was coupled to KLH, and antibodies generated in rabbits. Antibody has been affinity purified using control peptide-Sepharose.

Form & Storage

Control peptide Solution is provided in PBS, pH 7.4 at 1 mg/ml (100 ug/100 ul). Antiserum is supplied as neat serum (100 ul soln or lyophilized). Affinity pure antibodies were purified over the peptide-Sepharose column and supplied as 1 mg/ml soln in PBS, pH 7.4 and 0.1% BSA as stabilizer (100 ul in solution or Lyophilized).
The peptides and antibodies also contain 0.1% sodium azide as preservative. Lyophilized products should be reconstituted in 100 ul water and gently mixed for 15 min at room temp. All peptide/antibody

received in solution or

reconstituted from lyophilized vials should be stored frozen at -20oC or below in suitable aliquots. It is not recommended to store diluted solutions. Avoid repeated freeze and thaw.

Recommended Usage

Western Blotting (1:1K-5K for neat serum and 1-10 ug/ml for affinity pure antibody using ECL technique).

ELISA: Control peptide can be used to coat ELISA plates at 1 ug/ml and detected with antibodies (1:10-50K for neat serum and 0.5-1 ug/ml for affinity pure).

Histochemistry & Immunofluorescence: We recommend the use of affinity purified antibody at 1-20 ug/ml in paraformaldehyde fixed sections of tissues (1).

Specificity & Cross-reactivity

The 19 AA rat TR21-P control peptide is specific for rat TR2. It has no significant sequence homology with TR1 or gustducin or pheromone receptors. Antibody cross-reactivity in various species has not been studied. The TR21-P control peptide is available to confirm specificity of antibodies.

References:

1. Hoon MA et al (1999) Cell 96, 541-555; Lindemann B (1999) Nature Med. 5, 381-382

"Neat Antisera" are the unpurified antiserum and it is suitable for ELISA and Western.
"Affinity pure" antibodies have been over the antigen-affinity column and recommended for immunohistochemical applications.
"Control peptides" can not be used for Western as they are very short peptides. They are intended for ELISA or antibody competition studies.

List of Related Products

回頁首 回首頁

生工有限公司
MDBio, Inc.

全省免付費服務專線：0800-072-222 | 電話：(02)27474876 | 傳真：(02)28238024

網站：www.mdbio.com.tw | 電子郵件：mdbio@mdbio.com.tw

中華民國92年06月09日更新